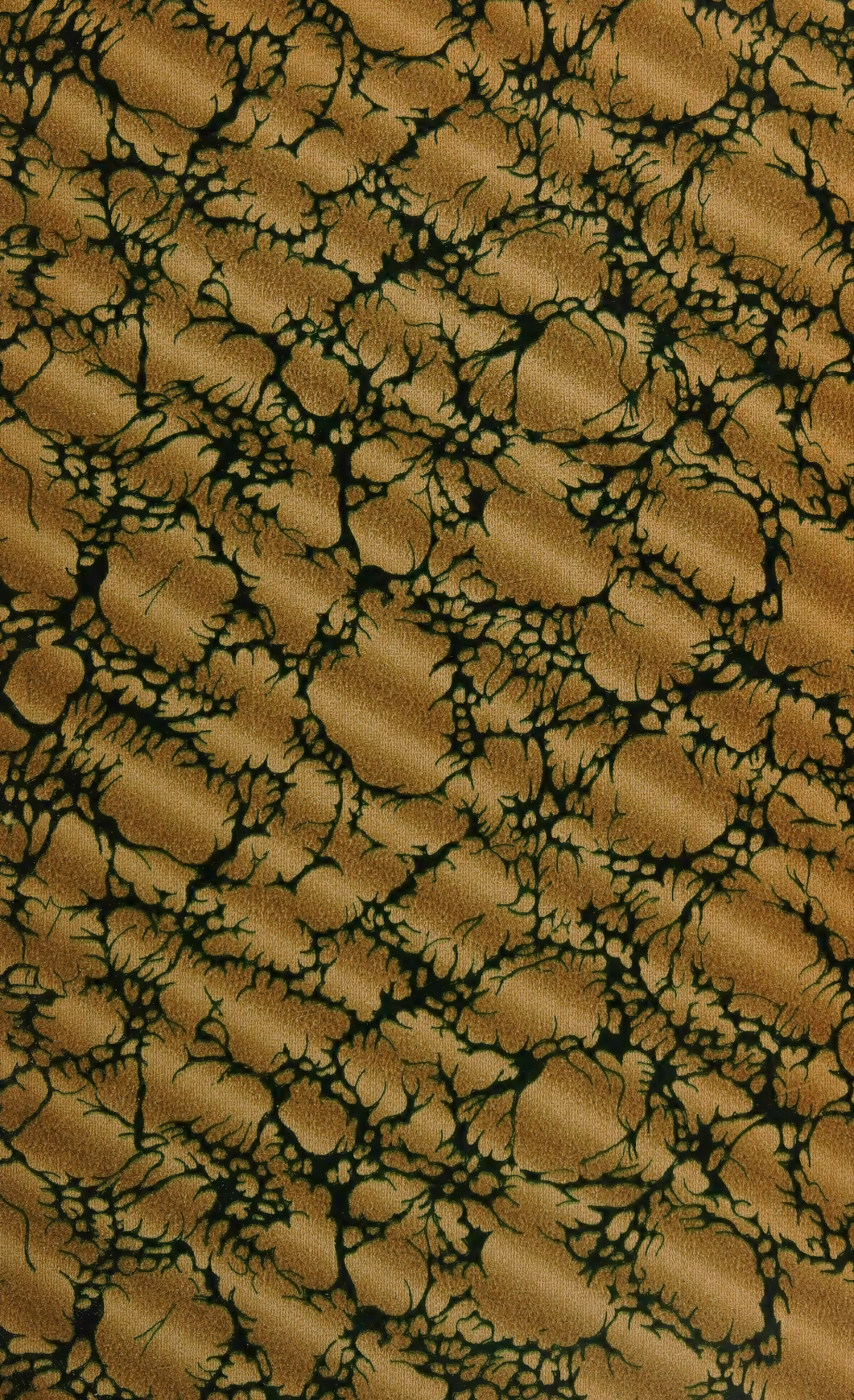
LIBRARY



LE CALCUL DES PROBABILITÉS,

SON ÉVOLUTION MATHÉMATIQUE ET PHILOSOPHIQUE

PAR

L.-Gustave Du PASQUIER

Diplômé de l'École polytechnique fédérale,
Docteur ès sciences,
Professeur de mathématiques supérieures
à l'Université de Neuchâtel.

PARIS
LIBRAIRIE SCIENTIFIQUE J. HERMANN
6, RUE DE LA SORBONNE, 6

1926

LE CALCUL
DES PROBABILITÉS,

SON ÉVOLUTION MATHÉMATIQUE ET PHILOSOPHIQUE

OUVRAGES DU MÊME AUTEUR

Le principe de la Relativité et les théories d'Einstein. Paris 1922. XVI + 511 p. avec 37 figures dans le texte.

Le développement de la notion de nombre. Tome III des *Mémoires de l'Université de Neuchâtel.* Neuchâtel et Paris 1921. VIII + 191 p.

Introduction à la science actuarielle. Neuchâtel et Paris 1918. VIII + 174 p.

Mathematische Theorie der Invaliditätsversicherung. Berne 1913. 153 p.

Sechs Vorträge über Astronomie. Lucerne 1911. VIII + 133 p.

Leonhardi Euleri Opera Omnia. Series 1, vol. 7. *Commentationes algebraicae ad theoriam combinationum et probabilitatum pertinentes.* Avec préface et analyse des travaux d'Euler sur les récréations mathématiques, le calcul des probabilités et ses applications à la statistique mathématique, à la théorie des erreurs d'observation et aux assurances sur la vie humaine. Leipzig et Berlin 1923. LVIII + 580 p. gr. in-4°. *Œuvres Complètes de Léonard Euler* éditées sous les auspices de la Société Helvétique des Sciences Naturelles.

LE CALCUL DES PROBABILITÉS,

SON ÉVOLUTION MATHÉMATIQUE ET PHILOSOPHIQUE

PAR

L.-Gustave Du PASQUIER

Diplômé de l'École polytechnique fédérale,
Docteur ès sciences,
Professeur de mathématiques supérieures
à l'Université de Neuchâtel.

PARIS
LIBRAIRIE SCIENTIFIQUE J. HERMANN
6, RUE DE LA SORBONNE, 6

1926

MATH-STAT

QA 273
D7
MATH.-
STAT.
LIBRARY

PRÉFACE

Le calcul des probabilités, né en France il y a plus de deux siècles et demi, est devenu l'une des branches les plus importantes des mathématiques appliquées. Et pourtant, malgré plusieurs essais qui ne manquent pas de valeur, il règne encore une confusion à peu près générale sur les fondements de cette théorie comme discipline mathématique, sur ses bases logiques et philosophiques. C'est d'autant plus étonnant qu'à l'époque actuelle, les questions axiomatiques éveillent le plus vif intérêt et que les applications du calcul des probabilités s'étendent de jour en jour davantage. Il est vrai que la découverte de Pascal et Fermat ne porta pas immédiatement ses fruits. L'histoire du calcul des probabilités est extrêmement capricieuse et l'on a pu dire avec raison que le hasard qui présida à sa naissance intervint maintes fois au cours de son développement. Cette histoire est non seulement très curieuse, mais aussi très complexe. Il faut y distinguer trois courants que d'ailleurs il est parfois difficile de démêler : 1° le courant mathématique ; il a progressé le plus rapidement ; 2° le courant pratique, orienté vers les applications statistiques et sociologiques ; 3° le courant philosophique. Grâce à l'action combinée de ces trois courants, il est actuellement possible de donner au calcul des probabilités une base mathématique inattaquable aux points de vue logique et philosophique. Exposer cette base nouvelle et l'évolution qui y aboutit est l'objet principal de ce livre. Quelques mots encore sur la voie suivie pour atteindre ce but.

Pour comprendre les interprétations, très différentes, données aux propositions de la théorie des chances et les objections soulevées contre elles, il est indispensable de connaître les théorèmes principaux du calcul des probabilités. Je devais donc, sous peine de perdre le fil de la déduction scientifique, commencer par exposer ces théorèmes, à moins de les supposer connus du lecteur ou de le renvoyer à d'autres ouvrages. J'ai préféré composer un livre où fût réuni tout ce qui est nécessaire à la compréhension pleine et entière du sujet. Après avoir esquissé l'évolution historique dans le chapitre premier, j'ai donc consacré les chapitres II, III et IV aux éléments du calcul des probabilités, au théorème de Jacques Bernoulli et à la théorie des écarts, en introduisant la définition habituelle de la probabilité mathématique basée sur la fameuse notion des cas également possibles.

Mon ouvrage aurait présenté une lacune, si je n'y avais pas mis un résumé de l'analyse combinatoire, la formule de Stirling et la connexion si importante entre le concept d'espérance mathématique et celui de valeur moyenne d'une variable dont la valeur dépend du hasard. Ceux qu'intéressent surtout les applications du calcul des probabilités à la statistique et à la biométrique trouveront les éléments nécessaires réunis dans cet ouvrage, y compris une table étendue de la transcendante de Kramp-Laplace, la fonction Thêta.

Pour arriver à saisir parfaitement la probabilité mathématique, le meilleur moyen est d'en suivre le développement historique. La célèbre loi biogénétique proclamée d'abord par les naturalistes : l'ontogenèse d'un individu est une répétition abrégée de la phylogenèse de sa race, se vérifie dans le domaine de l'intelligence. J'ai donc indiqué sommairement, dans les chapitres V, VI et VIII, les principaux stades par lesquels a passé la notion de probabilité mathématique dans le développement ininterrompu de l'esprit humain. Chemin

faisant, on voit se préciser de plus en plus la notion de hasard, inséparable de celle de probabilité.

A première vue, il semble paradoxal de parler de lois du hasard, alors que le hasard est l'antithèse de toute loi, la négation de toute règle. On verra quel sens il faut attribuer à ce terme de hasard pour éviter le conflit avec le déterminisme scientifique. Ainsi que le proclamait déjà Cournot, il y a moyen de concilier ces deux notions en apparence si opposées. La probabilité mathématique prend alors une valeur objective, aussi absolue qu'un théorème de géométrie par exemple, tout-à-fait indépendante de l'état plus ou moins avancé de nos connaissances, valeur qui subsisterait même pour une intelligence infinie et omnisciente.

Le chapitre V contient l'analyse des diverses interprétations du calcul des probabilités. Ce chapitre se termine par l'important principe des « zones de comportement », principe qui permet d'interpréter d'une manière nouvelle la loi des grands nombres et d'en mieux comprendre le sens.

Au chapitre VI, j'arrive à la notion du hasard dans le sens logique et par elle, grâce à la théorie des ensembles, à *une autre définition de la probabilité mathématique*. Elle éclaire d'un jour nouveau les fameux paradoxes du calcul des probabilités.

Avec la définition classique, il ne peut y avoir probabilité que là où il y a ignorance au moins partielle de certaines conditions ; la nouvelle définition, au contraire, est compatible avec le savoir le plus complet et la connaissance la plus entière de tous les détails. Je le prouve par plusieurs exemples.

Dans le chapitre VII sont traitées quelques applications importantes du calcul des probabilités à la physique théorique.

Dans le chapitre VIII vient la définition *stocastique* (*) de la

(*) Le terme de « stocastique » est dû à Jacques Bernoulli.

probabilité mathématique ; c'est une définition basée sur la théorie des ensembles collectifs. On pourrait appeler cette théorie, la plus récente de toutes, l'inversion des probabilités. Car en partant de la définition stocastique, on arrive à *démontrer* que, sous certaines hypothèses particulières, on peut statuer des « cas équipossibles » et que, ces hypothèses particulières admises, la probabilité mathématique est mesurée par le rapport du nombre des « cas favorables » à celui des « cas équipossibles ». Ce qui est définition dans la théorie classique devient ainsi un théorème. Ici encore, je ne suppose chez le lecteur aucune connaissance spéciale. Les notions nécessaires à l'exposé sont introduites au fur et à mesure des besoins, nettement définies, éclairées par des exemples.

Enfin, dans le chapitre IX, j'examine les rapports de la probabilité mathématique et de l'expérience. Je montre qu'avec la définition stocastique introduite au chapitre VIII, la connexion entre la théorie et l'expérience est tout-à-fait de même nature pour le calcul des probabilités que pour la géométrie, la mécanique rationnelle etc., tandis qu'il y a une différence de principe, si l'on adopte comme point de départ la définition classique ou celle introduite au chapitre VI. J'insiste aussi sur le postulat de la fortuité, en montrant les conséquences qu'entraînerait son abandon.

Voilà, esquissé à grands traits, le contenu de mon livre.

La plupart des traités de calcul des probabilités supposent des connaissances mathématiques étendues. On y utilise les symboles du calcul infinitésimal, et un lecteur non familiarisé avec l'analyse supérieure s'arrête bientôt découragé. J'ai voulu écrire un ouvrage qui fût à la portée de tous ceux dont les connaissances mathématiques ne dépassent pas l'algèbre élémentaire et ce qui fait partie de la culture générale commune aux hommes instruits. Dans la théorie des écarts seulement intervient un signe intégral ; c'est pour définir la fonction Thêta ; le lecteur pourra passer les formules qui lui

paraîtraient alourdir la marche de son étude et se contenter des résultats et de la table numérique. Dans les parties qui entrent en plein dans les mathématiques supérieures, j'ai préféré l'enseignement par l'exemple aux démonstrations abstraites. J'ai donné ces démonstrations là où elles sont faciles ou intuitives.

Si les origines du Calcul des probabilités sont modestes, la science nouvelle fondée par Pascal et Fermat a évolué d'une façon insoupçonnée. J'ai essayé d'exposer cette évolution grandiose sous une forme élémentaire et complète, du moins dans sa partie philosophique, poussant la théorie des écarts assez loin pour que le lecteur puisse aborder utilement les applications. J'ai recherché l'unité dans la conception de l'ensemble, l'ordre dans la disposition des parties, la simplicité et avant tout la clarté dans l'exposé du sujet.

L.-Gustave du Pasquier.

SOMMAIRE

	Pages	Numéros
Préface	V à IX	
Sommaire	XI	
Symboles et abréviations	XII	
Table des matières	XIII à XXI	
CHAPITRE PREMIER. — Introduction historique. 21 articles	1 à 30	1 à 21
CHAPITRE II. — Les éléments du calcul des probabilités. 28 articles	31 à 71	22 à 49
CHAPITRE III. — Les séries d'épreuves répétées. 26 articles	72 à 104	50 à 75
CHAPITRE IV. — Le théorème de Jacques Bernoulli et la théorie des écarts. 29 articles	105 à 151	76 à 104
CHAPITRE V. — Les diverses interprétations du calcul des probabilités. 25 articles	152 à 184	105 à 129
CHAPITRE VI. — La probabilité mathématique basée sur la théorie des ensembles ; la définition logique de la probabilité. 30 articles	185 à 224	130 à 159
CHAPITRE VII. — L'application du calcul des probabilités à la physique. 19 articles	225 à 247	160 à 178
CHAPITRE VIII. — La probabilité mathématique basée sur la théorie des collectifs ; la définition stocastique de la probabilité. 22 articles	248 à 279	179 à 200
CHAPITRE IX. — La probabilité mathématique et l'expérience. 12 articles	280 à 294	201 à 212
Appendice	295 à 304	
Deux tables de valeurs de la fonction transcendante de Kramp-Laplace, $\Theta(x)$:		
Table I à sept et à onze décimales, de $x = 0{,}00$ à $x = 4{,}80$ de centième en centième	296 à 298	
Table II à quatre décimales, de $x = 0{,}000$ à $x = 1{,}509$ de millième en millième	299 à 301	
Index des noms	302 à 304	

SYMBOLES ET ABRÉVIATIONS

$\equiv$ signifie « égal *par définition* à ».

$\neq$ signifie « différent de » ou « non égal à » (le signe de l'égalité, $=$, barré).

$\sim$ signifie « à peu près égal à » ou « semblable à ».

$<$ signifie « plus petit que » et $>$ « plus grand que ».

$\to$ signifie « tend vers » ou « tendant vers ».

$En(x)$ signifie « le plus grand nombre entier contenu dans x ».

$abs(x)$, ou simplement $as x$, signifie « valeur absolue de x ».

v. signifie voir.

n° signifie numéro ou paragraphe ou article.

Les formules sont numérotées à l'aide de deux nombres dont le premier indique le chapitre.

PLACE DES FIGURES :

Fig. 1, page 140 Fig. 2, page 141

Fig. 3, page 254 Fig. 4, page 257.

TABLE DES MATIÈRES

CHAPITRE PREMIER

INTRODUCTION HISTORIQUE

I. Les trois courants dans l'évolution du calcul des probabilités. — 1. Le courant mathématique. Le courant pratique orienté vers la sociologie. Le courant philosophique. Avance du courant mathématique. Analogie avec le développement du calcul infinitésimal. — 2. Difficulté de définir l'objet du calcul des probabilités ; ses causes. — 3. Plan de l'ouvrage.

II. L'invention du calcul des probabilités. Le courant mathématique. — 4. Blaise Pascal. Les deux questions posées par le chevalier de Méré. — 5. Pierre de Fermat. Ses mérites mathématiques. — 6. La méthode de Pascal. — 7. La méthode de Fermat. — 8. Les Anciens et le calcul des probabilités. — 9. Galilée et le jeu des dés. — 10. Le problème des « partis » jusqu'à Lagrange. — 11. Pierre-Simon de Laplace et sa *Théorie analytique des probabilités*. Son *Essai philosophique sur les probabilités*. — 12. L'évolution du « problème des partis ». — 13. La première phase de la théorie des chances : il s'agit de déterminer une valeur commerciale.

III. Le courant sociologique et la stocastique. — 14. John Graunt. Jean de Witt. Jean Hudde. Edmond Halley. — 15. L'*Ars coniectandi* de Jacques Bernoulli. Les recherches statistiques. — 16. La stocastique se crée.

IV. Le courant philosophique. — 17. La logique générale de l'incertain. La définition de la probabilité mathématique d'après Jacques Bernoulli. Elle devient la définition classique. Probabilités numériques et probabilités *anumériques*. — 18. *Le principe de la raison manquante* ou *principe de l'absence de raison justifiante*. Il est presque universellement adopté. Jacques Bernoulli, Laplace, Lourié. — 19. *Le principe de la raison péremptoire* formulé par Jean de Kries. — 20. La phase moderne du courant philosophique remonte à Cournot. Contrairement à ses prédécesseurs et contemporains, lui proclame la conciliabilité du

hasard et du déterminisme scientifique. — 21. La méthode axiomatique. Le « renversement » de la définition de la probabilité mathématique.

CHAPITRE II

LES ÉLÉMENTS DU CALCUL DES PROBABILITÉS

I. La définition classique de la probabilité mathématique. — 22. Etablissement de la formule fondamentale $p = f : s$. — 23 Définition de la probabilité mathématique. Evénements contraires. Application à quelques problèmes. — 24. Première observation : le cercle vicieux dans la définition classique. Réfutation. — 25 Deuxième observation : l'arbitraire dans la définition classique. Réfutation. — 26. Troisième observation : la difficulté dans la définition classique ; disjonction complète. Une erreur de LEIBNIZ. — 27. La restriction de « l'équipossibilité des cas ». Un faux raisonnement de D'ALEMBERT. — 28. Une application aux taux de mortalité ; l'urne de vie. — 29. Les quotients démographiques. Condition formelle que doit remplir une fraction « probabilitaire ». — 30. Certitude, probabilité infiniment petite, impossibilité. Les probabilités dites très grandes.

II. Les théorèmes fondamentaux. — 31. Méthodes directe et indirectes pour déterminer la probabilité mathématique Définitions préliminaires. — 32. *Le théorème d'addition* Conditions d'applicabilité. Probabilités partielle et totale. — 33 à 35. *Le théorème de multiplication.* Démonstration dans le cas d'événements indépendants. Répétition d'un même événement. Problème inverse. Applications. — 36. Un problème du jeu de trictrac. Notice historique sur la seconde question posée par le chevalier de Méré a Pascal. Généralisation au cas de 3 et de 4 dés. — 37. Démonstration du théorème de multiplication dans le cas d'événements dépendants. — 38. Le jeu de « pile ou face ». Application — 39. La loterie génoise.

III. L'analyse combinatoire. — 40. Les *permutations simples.* Démonstration de la formule. Exemple. — 41. *Les arrangements simples.* Démonstration de la formule. Exemple. — 42. *Les combinaisons simples.* Démonstration de la formule. Exemple. Introduction des factorielles $n!$ — 43. Le symbole d'Euler « n binomial p ». Propriétés des coefficients binomiaux. Le triangle arithmétique de Pascal. La formule du binôme de Newton. — 44. Les *permutations avec répétition.* Démonstration de la formule. — 45. Les *arrangements avec répétition.* Démonstration de la formule. Exemple : l'alphabet des télégraphistes. — 46. Les *combinaisons avec répétition.* Exemple. Démonstration de la formule.

IV. La formule de Stirling. — 47 à 49. Son but; sa nécessité; sa forme explicite. Applications numériques. L'erreur commise par son emploi. Son caractère asymptotique : l'erreur absolue $\to \infty$ avec n, mais l'erreur relative $\to 0$ quand $n \to \infty$. Exemples. Autres formules donnant une plus grande approximation.

CHAPITRE III

LES SÉRIES D'ÉPREUVES RÉPÉTÉES

I. La loi empirique du hasard ou loi des grands nombres. — 50 à 51. Cette loi est un fait expérimental et présente trois caractères distinctifs. L'exemple des pavés également mouillés. — 52. Définitions préliminaires. Fréquence absolue et fréquence relative. — 53. Variabilité des nombres de fréquence d'une série d'épreuves à une autre. — 54. Premier caractère de la loi empirique du hasard : le groupement des fréquences relatives autour d'un nombre central qui se trouve être égal à la probabilité mathématique dans le cas des jeux de hasard. Premier énoncé de la loi des grands nombres.

II. Les divers concepts d'écart. — 55. *L'écart relatif d'une série.* Définition. Exemple. Corollaires. Deuxième énoncé de la loi des grands nombres. — 56. *L'écart absolu* ou *la fluctuation d'une série.* — 57. L'unité d'écart absolu. L'unité d'écart relatif ou écart étalon. Définitions. Propriétés. — 58. *L'écart réduit* d'une série, λ. Définition. Sa grande importance provient de ce qu'il mesure à la fois l'écart absolu et l'écart relatif ; démonstration. Exemples. — 59 à 60. *La limite stocastique.* Différence avec la limite dans le sens de l'analyse mathématique. Troisième énoncé de la loi empirique du hasard. Remarque sur la connexion entre la théorie abstraite et la réalité concrète.

III. Les épreuves répétées et la formule du binôme de Newton. — 61. Analogie entre produits de binômes et probabilités composées. Exemples. — 62. Application au cas de deux événements exclusifs l'un de l'autre. — 63. Théorème général sur la signification « probabilitaire » de chaque terme dans le développement de $(p+q)^r$.

IV. La première partie du théorème de Jacques Bernoulli. — 64. Les deux problèmes fondamentaux qui se posent à propos d'une série d'épreuves répétées. — 65. Méthode générale de les résoudre. — 66 à 67. Application à un exemple numérique. — 68. *Résultats généraux* énoncés par anticipation. Asymétrie du développement lorsque $p \neq q$. Les cas exceptionnels de Markoff. — *Le résultat le plus probable d'une série d'épreuves identiques* : 69 à 70. Le nombre de fréquence le plus probable est déterminé univoquement ou bivoquement. Démonstra-

tion. Formules. — 71. Introduction de la fréquence relative dans les formules précédentes. *La première partie du théorème de Jacques Bernoulli.* Démonstration ; énoncés. Corollaire. — 72. Application à quelques problèmes.

La probabilité mathématique T'_0 du résultat le plus probable. — 73 à 75. Démonstration de la formule approchée $T_0 = 1 : \sqrt{2\pi pqr}$. Trois corollaires. Grandeur de l'erreur résultant de l'emploi de cette formule approchée. Autres expressions algébriques donnant T_0. Application.

CHAPITRE IV

LE THÉORÈME DE JACQUES BERNOULLI ET LA THÉORIE DES ÉCARTS

I. Probabilité mathématique que la fluctuation d'une série d'épreuves répétées prenne une valeur donnée d'avance. — 76 à 79. Probabilité mathématique $T(x)$ que l'écart absolu soit exactement égal à x. Propriétés de la fonction $T(x)$. Introduction de la « mesure de la précision absolue », h, et de la « mesure de la précision relative », h_r. Applications et exemples.

II. Probabilité mathématique que l'écart d'une série d'épreuves répétées reste compris entre des limites fixées d'avance. — 80 à 84. Introduction de la fonction Thêta. *La deuxième partie du théorème de Jacques Bernoulli.* Corollaires. Règle Exactitude de la formule de Laplace. Exemples ; application à un problème de Buffon. — 85. Le théorème d'Eggenberger. — 86. Applications au problème du nombre des épreuves nécessaires.

III. L'écart quadratique moyen d'une série d'épreuves répétées. — 87 à 88. Son importance. L'écart quadratique moyen de la fréquence absolue, μ ; l'écart quadratique moyen de la fréquence relative. Propriétés de ces grandeurs. Relations avec d'autres paramètres.

IV. L'écart probable d'une série d'épreuves répétées. — 89 à 90. Définitions. L'écart absolu probable. L'écart relatif probable. Modifications résultant du théorème d'Eggenberger. Application. Relations avec d'autres paramètres. L'écart probable est autre chose que « l'écart le plus probable ».

V. L'écart moyen d'une série d'épreuves répétées. — 91 à 92. Définitions. Théorèmes et corollaires. L'écart moyen de la fréquence absolue. L'écart moyen de la fréquence relative. Application. Relations avec d'autres paramètres. — 93. *Récapitulation* de ces 15 notions.

VI. Limites équiprobables dans les séries d'épreuves. — 94 à 95. Définitions. Exemples. Théorème général : leur indépendance du nom-

bre d'épreuves et de la probabilité élémentaire p. Tables. Applications.

VII. Les formules du hasard. — 96 à 97. Définitions. La première formule du hasard, $\mu^2 : e_a^2 = \pi : 2$. Les autres formules du hasard. Leur caractère singulier.

VIII. Sur la précision d'une série d'épreuves répétées. — 98 à 100. Comparaison de la théorie des écarts avec celle des erreurs d'observation. Les deux mesures de la précision d'une série d'épreuves. *La courbe des probabilités* ou courbe binomiale. Ses principales propriétés.

IX. Sur la dispersion dans les séries d'épreuves répétées. — 101 à 102. Définitions. Méthode. Application à 2 854 tirages de la loterie génoise.

X. Valeur moyenne et espérance mathématique. — 103 à 104. Définitions. La variable « casuelle ». L'espérance mathématique ; ses propriétés. Théorèmes. Exemple.

CHAPITRE V

LES DIVERSES INTERPRÉTATIONS DU CALCUL DES PROBABILITÉS

I. L'interprétation psychologique. — 105 à 106. Exposé. Les deux objections. Conclusion.

II. L'interprétation pratique. — 107. Exposé. Pourquoi elle est insoutenable.

III. L'interprétation logique. — 108 à 111. Exposé. Elle conduit au principe de l'absence de raison justifiante. Preuves que ce principe est insuffisant comme base de la théorie des chances. Exemples. Objection de la subjectivité non éliminée.

IV. L'interprétation empiriste. — 112 à 113. Exposé. Les nombres analogiques. Les principales objections. Conclusion.

V. L'interprétation inductive. — 114 à 116. Préliminaires sur l'induction Les principales objections contre cette interprétation. — 117. Circonstances *nomologiques* et *ontologiques*. — 118. Les « zones comportaires ». — 119. Possibilité objective et possibilité subjective.

VI. Le principe des zones de comportement. — 120 à 122. Définitions. Conditions de l'égale possibilité. Les zones comportaires doivent être *indifférentes, primaires, égales*. Les deux espèces de primordialité. — 123 à 126. Le jeu de la roulette rectiligne. Conditions d'applicabilité du principe des zones de comportement. Généralisations. — 127. Le principe des zones comportaires dans les jeux de hasard. — 128 à 129. *La loi des grands nombres expliquée par ce principe*. Corollaires.

CHAPITRE VI

LA PROBABILITÉ MATHÉMATIQUE BASÉE SUR LA THÉORIE DES ENSEMBLES. LA DÉFINITION LOGIQUE DE LA PROBABILITÉ

I. Considérations préliminaires. — 130 à 133. Choses complètes et choses incomplètes. Le principe de la participation. Réalité et existence. Une nouvelle interprétation du principe de l'absence de raison justifiante.

II. La probabilité mathématique et la théorie des ensembles. — 134 à 135. Rappel de quelques propriétés des ensembles. Le hasard dans le sens logique. — 136 à 138. La définition logique de la probabilité mathématique. Corollaires. Une remarque historique. — 139. *Probabilité et chance.* Poisson et Henri Poincaré.

III. Les paradoxes du calcul des probabilités. — 140 à 142. Le paradoxe de Bertrand; six solutions. Le rôle des ensembles dans la détermination de la probabilité mathématique.

IV. La probabilité mathématique et l'ignorance. — 143 à 146. D'après la nouvelle définition, l'ignorance même partielle n'est point nécessaire pour qu'il y ait probabilité mathématique assignable. Analogie avec les idées de Cournot. Application à l'astronomie (Laplace). — 147. Définition du *modèle* ou *échantillon d'un ensemble*. Propriétés. Exemples. — 148 à 149. Application heuristique du Calcul des probabilités. Gauss et les tables de logarithmes. — 150 à 153 Compatibilité parfaite de la probabilité mathématique avec le savoir complet et détaillé, grâce à la définition logique. Démonstration. Trois exemples (Gauss, Bertrand et Poincaré, Bruns ; la distribution des chiffres dans les tables de logarithmes).

V. La probabilité mathématique et le principe des zones de comportement. — 154 à 155. Ce principe est une source de probabilités, à côté de l'analogie et de l'induction. Sa grande portée. Son extension aux événements qui ne sont que plus ou moins semblables. Comparaison avec le principe de causalité.

VI. La probabilité mathématique et les régularités statistiques. — 156 à 157. L'interprétation de la loi des grands nombres dans le sens déterministe : que le hasard *doit* compenser, à la longue, les écarts qu'il produit, est erronée. Démonstration. Exemples. Quételet. Explication par le principe des zones de comportement. — 158. Nature apriorique de ce principe. — 159. Sur la nécessité des méthodes stocastiques « L'astronomie moléculaire ».

CHAPITRE VII

L'APPLICATION DU CALCUL DES PROBABILITÉS A LA PHYSIQUE

I. L'introduction du Calcul des probabilités dans la physico-chimie. — 160 à 161. Quand la réduction à une nécessité causale devient irréalisable, *possibilité* devient succédané pour *certitude*, et *probabilité* pour *vérité*. Constitution granulaire de la matière, de l'électricité (l'électron), du magnétisme (le magnéton de M. Pierre Weiss), de l'énergie (le quantum de M. Max Planck).

II. Résumé de la théorie cinétique des gaz. — 162 à 163. L'hypothèse cinétique fondamentale. Démonstration que *l'état le plus probable* est l'homogénéité. — 164 à 165. Inconvénients de la méthode stocastique. Les fluctuations. Pourquoi tout se passe comme si la loi approchée était rigoureusement exacte.

III. La répartition des vitesses moléculaires. La loi de Maxwell-Boltzmann. — 166 à 167. Impossibilité pour toutes les molécules d'avoir à chaque instant la même vitesse. La formule de Maxwell, ses trois formes. La répartition normale. Conséquences. Exemple numérique. On retrouve le principe des zones de comportement.

IV. Les phénomènes réversibles et le principe de Carnot. — 168 à 170. Objections à la théorie de Boltzmann Définition du phénomène idéalement réversible. Exemples. L'irréversibilité des phénomènes naturels. — 171. *Le principe de Carnot* ou deuxième principe de la thermodynamique. Corollaire. — 172. La formulation donnée par Clausius. L'entropie; elle ne diminue jamais.

V. Le déterminisme statistique et les fluctuations. — 173 à 175. Comment la théorie cinétique des gaz permet d'expliquer l'irréversibilité réelle constante des phénomènes physico-chimiques, et cela à l'aide de phénomènes idéalement réversibles L'influence théorique des fluctuations possibles. — 176. La différence de principe entre la thermodynamique classique et le point de vue moderne. — 177. L'évolution de l'Univers vers des états toujours plus probables. La formule de Boltzmann pour l'entropie. La méthode stocastique basée sur le principe des zones de comportement; sa portée philosophique.

VI. Autres applications du calcul des probabilités à la physique. — 178. Les phénomènes radioactifs. Le mouvement brownien et la formule d'Einstein. Les travaux de M. Paul Langevin. L'étude du rayonnement. La méthode stocastique dans une foule d'autres domaines, depuis l'astronomie à la biométrique et aux sciences sociales.

CHAPITRE VIII

LA PROBABILITÉ MATHÉMATIQUE BASÉE SUR LA THÉORIE DES COLLECTIFS. LA DÉFINITION STOCASTIQUE DE LA PROBABILITÉ

I. Le collectif empirique. — 179 à 181. Définitions préliminaires. Exemples de collectifs empiriques. — 182 à 184. La courbe de fréquence. La répartition normale. Décomposition d'une courbe en plusieurs courbes normales. Signification biologique. — 185 à 186. Les points de vue de QUÉTELET et de BRAVAIS. Exemples. Les collectifs asymétriques. Complications expérimentales.

II. Le collectif mathématique. La nouvelle définition de la probabilité mathématique. — 187 à 188. Quelques définitions. L'ensemble caractéristique. Exemples. — 189. Le postulat fondamental I ou *postulat de l'existence des valeurs limites*. La nouvelle définition de la probabilité mathématique. Sa supériorité sur la définition classique. Exemples. — 190. Corollaires. La diadose, fonction de fréquence ou loi de répartition du caractère distinctif parmi les éléments de l'ensemble considéré. — 191. Le postulat fondamental II ou *postulat de la fortuité*. Correspondance « arégulière ». Syllepte et collectif mathématique — 192 à 193. Corollaires. Impossibilité de la représentation explicite complète et détaillée. — 194. La fonction de fréquence comme fonction de point. Valeur moyenne.

III. Les quatre opérations simples pour déduire de nouveaux collectifs mathématiques. — 195 à 196 Le *choix admissible*. Le *mélangement*. Définitions. Exemples. Corollaires. — 197. Le renversement de la définition classique de la probabilité. — 198. Le *départagement*. Définition. Exemple. Corollaire : le théorème de Bayes. — 199. La *combinaison* de collectifs donnés. Définition. Exemples. Corollaire : le théorème de multiplication des probabilités. — 200. Combinaison et itération des quatre opérations simples. Objet du calcul des probabilités.

CHAPITRE IX

LA PROBABILITÉ MATHÉMATIQUE ET L'EXPÉRIENCE

1. Les deux postulats fondamentaux et l'expérience. — 201. Connexion entre notions abstraites et réalité physique en géométrie, en mécanique, en stocastique. — 202 à 203. Le postulat fondamental I ou postulat des valeurs limites suggéré par les faits observés. Exemples

de confirmation expérimentale. Analogies avec la géodésie. — 204. Le postulat fondamental II ou postulat de la fortuité suggéré par l'expérience. Vérifications diverses. Applications dans la science actuarielle.

II. **Le postulat de la fortuité et la probabilité mathématique.** — 205. Le postulat de la fortuité exclut du domaine des probabilités mathématiques une catégorie de problèmes traités dernièrement par les méthodes stocastiques. Exemple : les suites des $r^{\text{ièmes}}$ décimales dans une table de racines carrées ou de racines $n^{\text{ièmes}}$ des nombres entiers naturels. Trois théorèmes concernant la fréquence des divers chiffres, leur moyenne arithmétique, la fréquence relative des chiffres pairs et des impairs. — 206. Les théorèmes analogues de Franel pour les $r^{\text{ièmes}}$ décimales d'une table de logarithmes. — 207. Démonstration prouvant que le postulat de la fortuité n'est pas satisfait. La limite stocastique. — 208. L'abandon du postulat de la fortuité peut entraîner l'inexactitude du théorème des probabilités composées.

III. **Les systèmes logiques et l'expérience.** — 209 à 211. Les systèmes logiquement vrais. Les systèmes empiriquement vrais. Définitions. Propriétés. Exemples. — 212. L'application de théorèmes abstraits à la réalité physique. Cas de la géométrie, de la mécanique, du calcul des probabilités. Conclusion.

APPENDICE

Table I. — Valeurs de la fonction $\theta(x)$, à sept et à onze décimales, de centième en centième, pour l'intervalle de $x = 0{,}00$ à $x = 4{,}80$.

Table II. — Valeurs de la fonction $\vartheta(x)$, à quatre décimales, de millième en millième, pour l'intervalle de $x = 0{,}000$ à $x = 1{,}509$.

INDEX DES NOMS

LE CALCUL DES PROBABILITÉS SON ÉVOLUTION MATHÉMATIQUE ET PHILOSOPHIQUE

CHAPITRE PREMIER

INTRODUCTION HISTORIQUE

I. — *Les trois courants dans l'évolution du calcul des probabilités.*

1. — Blaise Pascal, dans une lettre de 1654, appelait le Calcul des probabilités, d'un nom aussi surprenant que mérité, « la géométrie du hasard ». Désigné autrefois sous des noms fort divers, ce calcul a cessé depuis longtemps d'être simplement une théorie des jeux de hasard. On a reconnu toute la valeur des méthodes si originales particulières à cette discipline ; leur emploi devient de plus en plus courant dans de nombreuses questions de physique, de biologie, voire même de psychologie, de sciences sociales et économiques. Bref, le Calcul des probabilités, né en France il y a plus de deux siècles et demi, est devenu de nos jours l'une des branches les plus importantes des mathématiques appliquées.

Trois courants se rencontrent dans la théorie des hasards : 1° un courant mathématique ; 2° un courant sociologique, orienté vers les applications pratiques ; 3° un courant philosophique. Nous aurons à parler successivement de ces trois courants. Disons d'emblée qu'ils n'ont pas progressé avec la même intensité. Le courant mathématique a pris une avance considérable. Les auteurs clas-

siques du calcul des probabilités ont bien touché à ses bases philosophiques, parfois même ont-ils cherché à les approfondir. Mais ce qu'ils en disaient dénotait encore une grande confusion, ne parvenait pas à réfuter péremptoirement les objections et laissait subsister des doutes fort légitimes sur la validité des résultats, surtout après les recherches de Condorcet (1), de Laplace (2) et de Poisson (3). Ces écrivains en effet ne craignirent pas d'appliquer, dans des travaux retentissants, toutes les ressources de l'analyse mathématique à la probabilité des témoignages et des jugements en matière criminelle et en matière civile.

Les philosophes, malgré de louables efforts, ne réussissaient pas non plus à lever ces doutes ni à écarter ces objections ; ils arrivaient au contraire à des conclusions souvent opposées et aux assertions les plus contradictoires. Dans les temps modernes seulement, les principes logiques et philosophiques de cette théorie ont été scrutés, précisés et développés d'une manière à peu près adéquate à son importance pratique.

Remarquons en passant que d'autres disciplines scientifiques eurent un sort semblable, par exemple le calcul infinitésimal inventé par Isaac Newton (1642-1727), puis par Gottfried W. Leibniz (1646-1716), indépendamment l'un de l'autre. Après la première phase, extraordinairement riche en découvertes et en applications imprévues, grâce surtout aux travaux des Bernoulli et du grand Euler, vint la phase métaphysique où l'on discuta sur la nature des infiniment petits, sur la légitimité du nouveau calcul et les limites de ses applications. Il n'y a pas très longtemps que les bases logiques et philosophiques du calcul différentiel et intégral

(1) Marie-Jean-Antoine-Nicolas-Caritat, marquis de Condorcet (1743-1794), *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Paris, 1785.

(2) Pierre Simon, comte de La Place (1749-1827), *Théorie de la probabilité des témoignages*. Ce travail est un chapitre de la *Théorie analytique des probabilités* à partir de la deuxième édition. Paris, 1814. V. n° 11.

(3) Siméon-Denis Poisson (1781-1840), mathématicien, physicien et astronome français, *Recherches sur la probabilité des jugemens en matière criminelle et en matière civile, précédées des règles générales du Calcul des probabilités*. Paris, 1837.

sont rigoureusement consolidées, indépendamment de toute métaphysique, et que l'analyse mathématique est entrée dans le troisième stade de son développement : la phase positive. On peut en dire autant de la théorie des chances. En 1848, Charles Gouraud (4) pouvait encore soutenir cette thèse à la Faculté des lettres de Paris : « Le Calcul des probabilités est une science dont les définitions sont encore à écrire, les principes à expliquer, les applications à restreindre et toute l'organisation à fonder. La partie mathématique seule mérite d'en être entièrement conservée ». Il n'en est plus de même aujourd'hui. Grâce à la théorie des ensembles et aux notions précises de mesure qu'elle a permis de formuler, il est possible actuellement d'asseoir le Calcul des probabilités sur des bases mathématiques et inébranlables aux points de vue logique et philosophique.

2. — La difficulté de définir l'objet du Calcul des probabilités exactement et brièvement à la fois, tient à sa très grande généralité. Il s'applique en effet aux choses les plus hétérogènes, concrètes ou immatérielles, et la connexion que l'on établit entre la mathématique et l'ensemble infini d'éléments aussi disparates ne peut pas être caractérisée en peu de mots. La difficulté est accrue par un double fait : 1° Les mots les plus employés dans le langage courant n'expriment en général pas des idées d'une précision mathématique, ou des notions rigoureusement définissables ni même univoquement déterminées. Des mots très fréquents tels que certain, incertain, fortuit, possible, probable, régulier, douteux, vrai, etc., ainsi que les substantifs correspondants : certitude, incertitude, hasard, chance, possibilité, probabilité, régularité, doute, présomption, conjecture, vérité, et quantité d'autres encore, ont une signification plus ou moins indécise et souvent plurivoque. Ces notions exercent néanmoins la plus grande influence sur nos décisions et nos actions journalières. Chacun attribue à ces mots un certain sens, sans pouvoir toujours l'énoncer clairement en paroles. C'est comme la monnaie courante : les pièces les plus

(4) Charles Gouraud, *Histoire du Calcul des probabilités, depuis ses origines jusqu'à nos jours*. Paris, 1848.

employées sont aussi les plus usées, à l'effigie peu nette ou effacée, aux contours souvent imprécis.

2° Les qualités ou les états d'âme exprimés par certains mots sont bien susceptibles de gradation, mais non de graduation ni de mesure, c'est-à-dire qu'on peut les comparer seulement d'une manière vague et *anumérique*. On peut dire par exemple qu'un fait A est un peu douteux, très douteux, extrêmement douteux etc. ; ou qu'il est un peu plus incertain qu'un autre fait B ou passablement plus incertain ou beaucoup plus incertain que lui ; mais dire qu'il est 3 1/2 fois plus incertain que B n'a pas de sens précis. De même, on peut dire d'un tableau qu'il est plus beau qu'un autre, beaucoup plus beau, infiniment plus beau ; mais dire qu'il est 3 1/2 fois plus beau n'a pas de sens assignable. Dans la probabilité mathématique, au contraire, on introduit des évaluations numériques : En lançant au hasard deux dés à jouer, il est 3 fois plus probable d'amener 7 points que d'en amener 11 ou que d'en amener 3 ; car la probabilité mathématique du premier événement est 6/36, celle d'amener 3 points, comme celle d'en amener 11, n'est que 2/36 (v. n° 26).

Que signifient au juste ces assertions ? Peut-on parler encore de probabilité mathématique lorsqu'il s'agit de la pluie et du beau temps ? Ou de l'issue incertaine d'une maladie ou d'une opération ? Ou de phénomènes sociaux ? Ces dernières probabilités sont-elles évaluables numériquement ? Si oui, sont-elles purement subjectives, variables d'une personne à l'autre ? Dépendent-elles de l'état de nos connaissances ? Dans quels cas une probabilité numériquement évaluée a-t-elle un sens assignable ? Il est nécessaire d'examiner à fond ces questions. On ne pourra fonder sur la notion de probabilité mathématique une science digne de ce nom que si ces probabilités ont quelque valeur objective indépendante de nos connaissances ou de notre ignorance ; de même qu'on pourrait difficilement ériger une Physique ou une Chimie où la définition des corps solides, des corps liquides, de l'oxygène etc., varierait suivant les lieux ou suivant l'état d'instruction du chimiste ou du physicien. Pour reprendre l'image ci-dessus, dans les recherches scientifiques, on ne doit utiliser que des monnaies à l'effigie très nette, aux contours très précis.

3. — Pour arriver à comprendre parfaitement la notion de probabilité mathématique, le meilleur moyen est d'en voir le développement historique. La célèbre loi biogénétique proclamée d'abord par des naturalistes : L'ontogenèse d'un individu est une répétition abrégée de la phylogenèse de sa race, se vérifie dans ce domaine de l'intelligence. Nous indiquerons donc, en les abrégeant, les principaux stades par lesquels l'idée de probabilité mathématique a passé, dans le développement ininterrompu de l'esprit humain. Nos investigations logiques et philosophiques aboutiront à une définition nouvelle de la probabilité mathématique, définition qui a l'avantage de concilier logiquement le déterminisme scientifique avec l'existence objective du hasard.

II. — *L'invention du calcul des probabilités. Le courant mathématique.*

4. — On a souvent, au début surtout, appelé le calcul des probabilités, « la théorie des hasards ». Et le hasard voulut qu'elle fût elle-même due à un hasard. Sur l'époque de l'invention de ce calcul, tout le monde est d'accord ; c'est le XVII^e^ siècle, « si fécond en grands hommes et en grandes découvertes qu'il est peut-être de tous les siècles celui qui fait le plus honneur à l'esprit humain » (Laplace). Et les nations qui se disputent si souvent, et parfois si âprement, l'honneur des plus grands progrès, n'embrouillent pas non plus la question : le berceau du Calcul des probabilités est en France. Pascal et Fermat se proposèrent et résolurent quelques problèmes qui exigeaient la connaissance des principes de ce calcul. A la satisfaction réciproque, ils virent qu'ils avaient trouvé la solution indépendamment l'un de l'autre. Tandis que les inventeurs du calcul infinitésimal, Newton et Leibniz, se sont laissés entraîner à une acerbe et peu glorieuse dispute de priorité, on voit les fondateurs du calcul des probabilités se réjouir ensemble de leur découverte commune et ne rien faire de ce qui pouvait appeler sur eux l'attention de la foule ou intéresser leur gloire.

Blaise Pascal (1623-1662) qui vécut tantôt à Clermont-Ferrand,

sa ville natale, tantôt à Rouen et à Paris, composa à seize ans le plus savant traité des coniques qu'on eût eu depuis l'antiquité. A dix-neuf ans, il inventa la machine à calculer ; à vingt-trois, il démontra les phénomènes de la pesanteur de l'air et détruisit une des plus grandes erreurs de l'ancienne physique. « Ayant achevé de parcourir le cercle des sciences humaines, il s'aperçut de leur néant et tourna ses pensées vers la religion. Depuis ce moment, toujours infirme et souffrant, il étonna le monde par ses *Lettres provinciales* où il donna le modèle de la plus parfaite plaisanterie comme du raisonnement le plus fort et où il fixa la langue que parlèrent Bossuet et Racine ; dans les courts intervalles de ses maux, il résolut, par distraction, un des plus hauts problèmes de la géométrie et jeta sur le papier des *Pensées* qui tiennent autant du dieu que de l'homme » (Châteaubriand). A cet austère janséniste, le chevalier de Méré, homme d'esprit autant que joueur passionné, vint proposer, en 1654, deux difficultés de jeu dont ni la réflexion ni l'expérience n'avaient pu lui procurer la solution.

α) Voici la première question : En combien de coups peut-on parier avec avantage d'amener le *sonnez* au jeu de trictrac, c'est-à-dire le double six en jetant simultanément deux dés ordinaires (*). Nous traiterons ce problème au chapitre II, n° 36.

β) La deuxième question est devenue célèbre sous le nom de *problème des partis* : Deux joueurs *A* et *B* également habiles, c'est-à-dire ayant, à chaque partie, égale chance de gain ou de perte, ont mis chacun un enjeu convenu, disons 8 francs pour fixer les idées. Le gagnant aura droit à la totalité de la mise, donc à 16 francs. Chaque partie jouée fait gagner un point, soit à l'un soit à l'autre des deux joueurs. Le gagnant sera le premier qui obtiendra 3 points. Pour terminer le jeu, il faut donc faire au moins 3 et au plus 5 parties. Or, après la première partie, les joueurs sont obligés de se séparer, on le suppose, sans attendre la décision du sort.

(*) On dit que *sonnez* provient du dicton des joueurs de trictrac, s'écriant quand ils amenaient les deux 6 : « Sonnez, le diable est mort ». Mais l'étymologie la plus probable est celle qui fait dériver *sonnez*, dont l'ancienne forme est *senes*, *sines*, *sannes*, du latin *seni* = « qui sont six » (Littré, *Dictionnaire de la langue française*, article *sonnez*).

Quel est *le parti* de chaque joueur? En d'autres termes : dans quelle proportion doivent-ils répartir « équitablement » l'enjeu de 16 francs? Quelle est la répartition du fonds commun des mises exactement proportionnelle à la valeur des droits que le hasard de la partie leur avait jusque-là assurés, et à celle des espérances que l'état présent du jeu leur permettait de concevoir encore? Il s'agissait donc de mesurer numériquement le degré de vraisemblance dont pouvaient être dignes de simples conjectures. L'histoire offre ici un exemple remarquable que l'occasion la plus imprévue peut susciter la découverte la plus extraordinaire. Aucun précédent n'autorisait de penser qu'il fût possible d'employer l'arithmétique à un usage aussi hardi. « Pascal en fit l'épreuve et trouva dans son génie le secret d'y réussir. Cette démonstration étonnante réveillait soudain dans l'intelligence de l'homme un ressort inconnu dont personne encore ne pouvait prévoir ni toute la portée, ni toute la puissance » (Gouraud (4), p. 5; voir n° 10 et n° 12).

5. — Pierre de Fermat (1601-1665) fit ses études de droit et devint membre du parlement de Toulouse (1631), puis fut ennobli quelques années plus tard. Il jouissait d'une réputation méritée de juriste perspicace et de sévère intégrité. Au milieu des devoirs absorbants de sa charge, il sut trouver le temps de s'occuper de littérature et de science. D'une érudition aussi vaste que profonde, il fut l'un des plus grands inventeurs que les mathématiques aient jamais eus. Sa correspondance prouve qu'il connaissait la notion et l'emploi des coordonnées, fondement de la géométrie analytique qui a fait la gloire de Descartes. D'Alembert, Lagrange, Laplace, attribuent à Fermat la gloire de la première idée du calcul infinitésimal, si débattue depuis entre l'Allemagne et l'Angleterre. Dans la théorie des nombres, Fermat a fait revivre l'arithmétique supérieure ébauchée par Diophante. On doit à Fermat de nombreux théorèmes d'arithmomie ; l'un en particulier, appelé le grand théorème de Fermat dont la démonstration est restée dans son esprit, a acquis au commencement de notre siècle, après avoir fatigué sans succès les Euler, les Lagrange, les Gauss, une popularité exagérée

grâce au legs Wolfskehl (*). Enfin, Fermat partage avec Pascal l'honneur de l'invention du calcul des probabilités. On le voit, le souvenir du génie de Fermat est associé aux plus grands travaux de son siècle. Pascal le nomme le premier homme du monde et avoue ne pas toujours pouvoir le suivre dans ses recherches. « Cherchez ailleurs », lui écrit-il, « qui vous suive dans vos inventions numériques. Pour moi, je vous confesse que cela me passe de bien loin ; je ne suis capable que de vous admirer ».

Unis par une confraternité de goût et de génie, Pascal et Fermat, les plus grands géomètres de leur époque, étaient en correspondance mathématique suivie. Il arrivait rarement à l'un de faire quelque découverte sans en conférer avec l'autre. Fidèle à cette coutume, Pascal envoya à Fermat l'énoncé des problèmes du chevalier de Méré et les solutions qu'il en donnait. « Demander l'avis d'un pareil inventeur sur le mérite d'une découverte, était lui fournir au moins l'occasion de la perfectionner. Ce fut en effet ce qui arriva » (Gouraud (4), p. 7). Pascal s'en était tenu strictement aux deux problèmes posés. Sa méthode de résolution, ravissante par son ingéniosité autant que par son originalité, semblait inapplicable au cas d'un plus grand nombre de joueurs. Fermat trouva une méthode basée sur la théorie des combinaisons, d'où découle une règle générale qui permet de partager équitablement l'enjeu non seulement entre deux, mais entre un nombre quelconque de joueurs et quel que soit le moment où il conviendrait d'interrompre le jeu. Appliquée au problème particulier du chevalier de Méré, la méthode de Fermat conduit aux résultats donnés par Pascal. Le perfectionnement de Fermat complétait la découverte de Pascal : le calcul des probabilités était doté de sa première méthode générale. Dans le premier stade du développement, on voit en effet les progrès du calcul des probabilités marcher de pair avec ceux de l'analyse combinatoire, et les premiers écrits concernant la théorie des chances sont effectivement aussi les premiers sur l'analyse combinatoire.

(*) Cent mille marks de récompense à celui qui démontrera que l'équation

$$x^n + y^n = z^n$$

n'admet aucune solution *en nombres entiers* non nuls, dès que n représente un nombre entier supérieur à 2.

6. — L'ingénieux raisonnement de Pascal pour résoudre le problème des partis consiste à supposer que le jeu se continue encore d'un coup et à trouver de quelle partie de l'enjeu ce coup détermine le sort. Pascal résolut le problème du chevalier de Méré en trois pas successifs (le gagnant étant par hypothèse celui qui le premier obtient 3 points).

Premier pas. — 3 parties ont été jouées ; *A* a déjà gagné 2 points et *B* 1 point ; comment partager l'enjeu de 16 francs ? Supposons qu'on joue encore une partie ; si *A* la gagne, il aura 3 points et touchera les 16 francs, *B* ne touchera rien ; si *B* la gagne, chaque joueur aura 2 points ; c'est l'égalité de chance et il est juste de donner à chacun la moitié de l'enjeu, soit 8 francs. En résumé : si *A* gagne la partie en question, il a droit à 16 francs ; si *A* la perd, il a droit à 8 francs ; différence 8 francs. Si donc les joueurs conviennent de se séparer sans terminer le jeu, *A* pourra tenir à son partenaire *B* le raisonnement suivant : « Que je gagne ou que je perde, j'aurai en tout cas 8 francs ; quant au reste de mon gain possible, 8 francs, je pourrais le gagner aussi bien que vous, nos chances sont identiques, partageons-le donc en parts égales ; vous toucherez 4 francs, et moi 4 francs plus les 8 qui me reviennent en tout état de cause ». Ainsi, dans cette première possibilité, *A* reçoit 12 francs $= \frac{3}{4}$ de la mise totale, *B* touche le reste $= \frac{1}{4}$ de la mise.

Deuxième pas. — 2 parties ont été jouées ; *A* a gagné 2 points et *B* zéro ; comment partager l'enjeu ? Supposons qu'on joue encore une partie ; si *A* la gagne, il aura 3 points et touchera les 16 francs, *B* ne touchera rien ; si *B* la gagne, il aura 1 point et *A* 2 points ; c'est la situation qu'on vient d'examiner, il revient alors 12 francs à *A* et 4 à *B*. En résumé : si *A* gagne la partie en question, il a droit à 16 francs ; si *A* la perd, il a droit à 12 francs ; différence 4 francs. Si donc les joueurs conviennent de se séparer sans terminer le jeu, *A* pourra tenir à son partenaire *B* le raisonnement suivant : « Que je gagne ou que je perde, j'aurai en tout cas 12 francs ; quant au reste de mon gain possible, 4 francs, je

pourrais le gagner aussi bien que vous, nos chances sont identiques, partageons-le donc en deux parts égales ; vous toucherez 2 francs, et moi 2 francs plus les 12 qui me reviennent en tout état de cause ». Ainsi, dans cette deuxième possibilité, *A* reçoit 14 francs $= \frac{7}{8}$ de la mise totale ; *B* touche le reste $= \frac{1}{8}$ de la mise.

Troisième pas. — Une seule partie a été jouée ; *A* a gagné 1 point et *B* zéro point. C'était le problème posé par de Méré : Comment partager l'enjeu dans ce cas ? Supposons, dit Pascal, qu'on joue encore une partie. Si *A* la gagne, il aura 2 points contre zéro ; c'est la situation qu'on vient d'examiner : il revient alors 14 francs à *A* et 2 à *B* ; si au contraire *B* la gagne, il aura 1 point contre 1 ; c'est l'égalité de chance et il est juste de donner à chacun la moitié de l'enjeu, soit 8 francs. En résumé : si *A* gagne la partie en question, il a droit à 14 francs ; si *A* la perd, il a droit à 8 francs ; différence 6 francs. Si donc les joueurs conviennent de se séparer sans terminer le jeu, *A* pourra tenir à son partenaire *B* le raisonnement suivant : « Que je gagne ou que je perde, j'aurai en tout cas 8 francs ; quant au reste de mon gain possible, 6 francs, je pourrais le gagner aussi bien que vous, nos chances sont identiques, partageons-le donc en deux parts égales. Je toucherai ainsi 3 francs plus les 8 francs qui me reviennent en tout état de cause. » Dans ce cas donc, *A* touche 11 francs $= 11/16$ de la mise totale ; *B* reçoit le reste $= 5/16$ de la mise.

Tel était le résultat de l'ingénieuse « méthode des partis » imaginée par Pascal : quand deux joueurs également habiles sont à 1 contre zéro, 3 étant le maximum, il revient au premier $11/16$, au second $5/16$ de l'enjeu.

7. — La méthode imaginée par Fermat pour résoudre le problème des partis est basée sur la remarque que, s'il manque au premier joueur encore m et au second encore n points, le jeu se terminera sûrement après $m + n - 1$ parties. Dans le problème du chevalier de Méré, il manquait à *A* 2 points, à *B* encore 3 points ; le jeu se terminait donc sûrement au bout de $3 + 2 - 1 = 4$ parties. Elles donnent lieu aux 16 combinaisons suivantes où la lettre a désigne

une partie gagnée par le joueur A, et la lettre b une partie gagnée par B.

aaaa	*aabb*	*baba*	*bbab*
aaab	*abab*	*bbaa*	*bbba*
aaba	*abba*	*abbb*	*bbbb*
abaa	*baab*	*babb*	
baaa			

De ces 16 combinaisons, toutes équipossibles par hypothèse, les 11 premières sont favorables à A et les 5 dernières à B, puisqu'il est alors le premier à avoir 3 points. Il est dès lors juste de partager l'enjeu dans le rapport de 11 à 5, c'est-à-dire d'attribuer à A les 11/16 et à B les 5/16 de l'enjeu. Pascal, enchanté de voir ses résultats confirmés par Fermat, lui écrivit : « Je ne doute plus maintenant que je ne sois dans la vérité, après la rencontre admirable où je me trouve avec vous... Je vois bien que la vérité est la même à Toulouse qu'à Paris ». Pascal étendit la méthode de Fermat au cas de trois joueurs.

8. — Les Anciens paraissent avoir ignoré entièrement le calcul des probabilités. L'érudition moderne en a cependant trouvé quelques traces antérieures au temps de Pascal : dans un poème en latin barbare intitulé *De Vetula*, dans un commentaire (Venise, 1477) de la *Divina Commedia* (5) de Dante Alighieri (1265-1321) et dans les écrits de plusieurs mathématiciens italiens du moyen âge et de la Renaissance : Hieronymo Cardano (1501-1576) (6), Luca Paciuolo (1445-1514) (7), Peverone (8), Nicolo Tartaglia (1500-1557) (9). « Mais ces essais grossiers, d'une analyse extrêmement

(5) Ce commentaire, plus intéressant pour l'étymologiste que pour le mathématicien, se rapporte à la première ligne du sixième chant du *Purgatorio*, à propos du mot *Azari*. Voir Encyclop. sc. math., t. I, vol. 4, fasc. 1, p. 2-3, note 6.

(6) *Œuvres de Cardan*, Lyon, 1663, t. I, p. 262-276.

(7) *Summa de Arithmetica, Geometria, Proportioni et proportionalita*, traité qui eut assez de vogue en son temps, écrit en 1487 à Perugia, imprimé en 1494 à Venise ; chap. X : « Tractatus De straordinariis », fol. 197.

(8) Cité d'après Gouraud (4), p. 3.

(9) *General Trattato di numeri et misure*, 1556-1560, Parte 2.

fautive et restés tous également stériles, ne sont dignes des regards ni de la critique, ni de l'histoire. Pascal en ignora jusqu'à l'existence, et quand il les eût connus, ils n'auraient pu lui servir ni de guides, ni de modèles ». On ne peut qu'approuver la dernière partie de la dernière phrase, même si l'on trouve le jugement de Gouraud un peu trop sévère ; en tous cas, il est injuste en oubliant Galilée.

Galileo Galilei (1564-1642), le grand savant de Florence, fondateur de la Mécanique classique et inventeur du télescope, sa vie si riche en travaux, en luttes et en polémiques scientifiques, ses multiples tribulations, son procès avec le Saint Office de l'Inquisition, sa condamnation et son abjuration sensationnelle de la doctrine de Copernic, sa fin triste d'homme aveugle toujours sous la surveillance de l'Inquisition, tout cela est universellement connu.

9. — Parmi les manuscrits laissés par Galilée, on a trouvé un écrit intitulé *Considerazione sopra il Giuoco dei Dadi* (10). Ces « Considérations sur le jeu des dés » se rapportent au jeu de passe-dix : Un joueur jette au hasard 3 dés ordinaires ; si la somme des points amenés est supérieure à 10, il gagne ; si la dite somme est inférieure ou égale à 10, il perd. Un amateur de ce jeu très en vogue à cette époque, ami de Galilée, ayant observé les coups et discuté les chances, remarqua qu'il gagnait plus souvent par le point 11 que par 12 ; il s'étonnait aussi de voir que 10 arrivait plus souvent que 9, car, disait-il, avec 3 dés, il y a 6 manières d'amener 9 et 6 manières d'amener 10 points, comme il y a 6 manières d'amener 11 et 6 façons d'amener 12 ; et pourtant, on compte 108 fois le point 11 contre 100 fois le point 12. Il soumit cette contradiction apparente à Galilée. Celui-ci donna l'explication en établissant mieux le compte et faisant remarquer que les 6 manières ne sont pas équivalentes, parce qu'inégalement possibles. Au lieu de 6 chances, il montre distinctement 27 chances pour 11 points et

(10) *Opere di Galileo Galilei*, Firenze, 1855, vol. 14, p. 293-296. Ed. nazionale, Firenze, 1898, VIII, p. 591-594. Cette pièce, antérieure à 1642, fut publiée pour la première fois à Florence, dans l'édition de 1718 des *Œuvres de Galilée*, t. III, p. 119-121.

25 seulement pour 12 points, ce qui donne bien le rapport de 108 à 100. Cela ressort du tableau suivant.

9 points		10 points		11 points		12 points	
amenés par	Nombre des cas possibles	amenés par	Nombre des cas possibles	amenés par	Nombre des cas possibles	amenés par	Nombre des cas possibles
1 + 2 + 6	6	1 + 3 + 6	6	1 + 4 + 6	6	1 + 5 + 6	6
1 + 3 + 5	6	1 + 4 + 5	6	1 + 5 + 5	3	2 + 4 + 6	6
1 + 4 + 4	3	2 + 2 + 6	3	2 + 3 + 6	6	2 + 5 + 5	3
2 + 2 + 5	3	2 + 3 + 5	6	2 + 4 + 5	6	3 + 3 + 6	3
2 + 3 + 4	6	2 + 4 + 4	3	3 + 3 + 5	3	3 + 4 + 5	6
3 + 3 + 3	1	3 + 3 + 4	3	3 + 4 + 4	3	4 + 4 + 4	1
	25		27		27		25

En effet, 4 + 4 + 4 n'est pas équivalent à 1 + 5 + 6 ni à 2 + 5 + 5, quoique les trois possibilités donnent 12 points comme somme. La première combinaison est unique : chacun des trois dés doit amener 4. La deuxième possibilité, au contraire, représente 6 combinaisons, par la même raison qu'avec 3 lettres distinctes, par exemple *a*, *b*, *r*, on peut former 6 mots différents. La troisième possibilité, 2 + 5 + 5, représente 3 combinaisons, par la même raison qu'avec 3 lettres dont 2 égales, par exemple *a*, *a*, *r*, on ne peut former que 3 mots différents.

Remarque I. — On voit facilement qu'il y a chance égale de perdre ou de gagner au jeu de passe-dix. Dans les dés ordinaires, les faces opposées portent toujours les nombres 1 et 6 ; 2 et 5 ; 3 et 4. La somme des points de deux faces opposées est donc toujours = 7. Si l'on appelle « inférieure » la face sur laquelle repose le dé et « supérieure » celle qui est opposée à la face inférieure, on

voit que, avec 3 dés, la somme des points des 3 faces inférieures et des 3 faces supérieures donne $3.7 = 21$ au total. Dès lors, si la somme des points des 3 faces supérieures est plus grande que 10, la somme des points des 3 faces inférieures sera moindre que 11. On en conclut bien l'égalité de chance d'amener, avec 3 dés, plus de 10 points ou moins de 11 points, c'est-à-dire de gagner ou de perdre au jeu de passe-dix.

Remarque II. — Il va de soi que bien avant Pascal et Fermat, les hommes ont basé leur conduite et leurs actions sur des probabilités, tout comme on a fait de la géométrie avant que ne parussent les *Eléments* d'Euclide. Néanmoins, on peut dire avec Gouraud que c'est grâce à Pascal qu'une science « sans racine dans le passé et du plus brillant avenir, a marqué son rang dans le système agrandi des connaissances humaines ». Ce n'est en effet qu'à partir de ce moment, où pour la première fois on exprima sous forme mathématique certaines notions et certains principes, que le calcul des probabilités put devenir une discipline mathématique. L'impulsion, une fois donnée, a continué d'agir.

10. — *Le problème des partis* ou « problem of points » des Anglais, « Teilungsproblem » des Allemands, est, avec le problème dit de la ruine des joueurs ou de la durée du jeu, celui qui a le plus fait mettre en œuvre toutes les ressources de l'analyse mathématique supérieure. Les plus grands géomètres y ont essayé la puissance de leurs méthodes, depuis les Bernoulli, Lagrange, Laplace jusqu'aux auteurs modernes, et il n'est pas sans intérêt de jeter un coup d'œil sur l'évolution de ce problème particulier, origine du calcul des probabilités.

Résoudre le problème des partis revient à déterminer la probabilité pour chaque joueur de gagner la partie. Pascal n'a pas cessé de perfectionner sa « méthode des partis » et lui a consacré de nombreuses pages de son célèbre *Traité du triangle arithmétique* (11). Il est remarquable que l'idée fondamentale de cette méthode ait conduit Lagrange un siècle plus tard à une solution

(11) Ecrit en 1654, mais publié en 1665 seulement, trois ans après la mort de Pascal. V. *Œuvres de Blaise Pascal*, Paris, 1819, t. V.

générale inattendue du problème des partis. Un petit traité de Huygens (1629-1695), paru en 1656 (12), constitue le premier essai de rassembler les nouveaux principes en un système et de combler une lacune laissée par Pascal et Fermat. Huygens résout à son tour le problème des partis pour les cas de 2 et de 3 joueurs, mais sans faire usage de nouvelles méthodes. Il en est de même de Jacques Bernoulli (1654-1705) dont le nom est indissolublement lié à l'histoire du calcul des probabilités, grâce à son ouvrage monumental sur « l'Art de conjecturer », *Ars conjectandi* (13). Tous ces auteurs s'en tiennent à l'hypothèse que les chances de gagner ou de perdre une partie isolée, ou 1 point, sont égales pour tous les joueurs. Abraham de Moivre (1667-1754), huguenot réfugié en Angleterre en 1685, après la révocation de l'édit de Nantes, reçu membre de la Société royale de Londres en 1697, fut le premier qui traita le cas où ces probabilités de gagner 1 point sont inégales pour les différents joueurs (14). C'était un pas dans la voie de la généralisation.

Pierre Rémond de Montmort (1678-1719), pendant quelque temps chanoine de Notre-Dame de Paris, dès 1716 membre libre de l'Académie des sciences de Paris, a publié la première formule générale résolvant le problème des partis, pour le cas de 2 joueurs dont l'un a la probabilité p de gagner 1 point, l'autre la probabilité $q = 1 - p$. Sa méthode est le développement de celle de Fermat (15).

(12) Christian Huygens (Hugenius), *De ratiociniis in ludo aleae* (Des partages dans le jeu de dé) ; écrit d'abord en hollandais, ce travail fut traduit en latin par Franciscus van Schooten (1615-1660) qui le publia comme appendice de ses *Exercitationum mathematicarum Libri quinque*, Lugdun. Batav., 1657.

(13) Voir la note 25, p. 22.

(14) *De Mensura sortis, seu de probabilitate eventuum in ludis...*, Philos. Transact. vol. XXVII, p. 213-264, Lond. 1711. Ce mémoire fut développé par de Moivre qui en fit un livre : *The Doctrine of Chances*, etc. prem. éd. Londres, 1718 ; 2e éd. 1738 ; 3e éd. 1756. Dans les seconde et 3e éditions, de Moivre indique une règle pratique pour résoudre le problème des partis dans le cas d'un nombre quelconque de joueurs.

(15) *Essay d'Analyse sur les Jeux de Hazards*, 2e éd., p. 232-248, Paris, 1713. Dans la première édition de cet ouvrage, Paris, 1708, de Montmort n'avait envisagé que le cas où les joueurs ont même habileté.

L'illustre mathématicien français Joseph-Louis Lagrange (1726-1813), né à Turin où il devint professeur à l'Ecole d'artillerie et qui, appelé par le roi de Prusse Frédéric-le-Grand, accepta, en 1766, de succéder au grand Euler comme Directeur de la Classe de mathématiques de l'Académie des sciences de Berlin, pour venir finalement habiter Paris, résolut à son tour et complètement le problème des partis, pour un nombre quelconque de joueurs, à l'aide d'une méthode toute nouvelle faisant usage des équations aux différences finies (16).

11. — Pierre Simon comte de La Place (1749-1827) qui, fils d'un cultivateur pauvre, parvint au faîte des honneurs, s'est occupé plusieurs fois du problème des partis : d'abord dans son *Mémoire sur la Probabilité des Causes par les Evénemens*, Paris, 1774, puis dans ses *Recherches sur l'intégration des équations différentielles aux différences finies*, etc. Paris, 1776, enfin dans sa célèbre *Théorie analytique des Probabilités*, dont la première édition, dédiée à Napoléon-le-Grand, parut en 1812 à Paris. Il y donne la solution générale pour le cas d'un nombre quelconque de joueurs.

Il faut dire que dans ce grand ouvrage, Laplace a fondu en un tout presque tous les mémoires qu'il avait écrits auparavant sur la théorie des chances. La *Théorie analytique des probabilités* a donné à la science mathématique des hasards cette unité qui lui manquait encore après un siècle et demi d'existence. La Place a su réunir dans cette œuvre les multiples procédés usités jusqu'alors, grâce au calcul des « fonctions génératrices » imaginé par lui en 1809 déjà, méthode difficile, abordable seulement au mathématicien, mais d'une souplesse et d'une étendue prodigieuses. La *Théorie analytique des probabilités* a fait époque. On peut même dire que depuis lors, en tant qu'il s'agit des développements mathématiques de ce calcul, rien d'essentiellement nouveau n'y a été ajouté ; seul son domaine d'application, comme il fallait s'y attendre, a continué à s'élargir et à se préciser.

(16) Voir *Recherches sur les suites récurrentes... et sur l'usage de ces équations dans la théorie des hazards*. Nouv. Mém. Acad. Berlin pour 1775, publ. 1777, p. 183-272 ; en particulier, p. 240-272.

Il convient d'ajouter que la *Théorie analytique des probabilités*, depuis sa deuxième édition, Paris, 1814, est précédée d'une *Introduction* qui a paru séparément sous le titre *Essai philosophique sur les probabilités*, développement d'une leçon que Laplace avait donnée en 1795, lors de la création de l'Ecole normale supérieure. « L'auteur, sans le secours de l'analyse, y expose les grands principes et les grands résultats de la Théorie des hasards et montre qu'elle s'applique non seulement à des problèmes de pure curiosité, mais aux questions les plus intéressantes de la vie des individus et des nations » (Gouraud (4), p. 113). Le succès de cet *Essai philosophique* fut très grand (*) et une autorité considérable qui n'a fait que s'accroître, s'attacha dès l'origine à cette publication.

Par son *Essai philosophique sur les probabilités*, Laplace a fortement influé sur le courant philosophique dont nous aurons à parler (v. nos 18, 21 et le chap. V). Mais ses profondes études ont aussi communiqué au courant sociologique une impulsion qui n'a cessé d'aller en grandissant. On peut même dire que le commencement du XIXe siècle est une de ces époques d'engouement comparables à celle qui suivit la publication de l'*Ars conjectandi* (v. no 15). De toutes parts on s'efforçait de réaliser le plan tracé par Laplace. On appliqua le calcul des probabilités à l'économie politique, à la médecine, enfin aux sciences sociales. On peut aussi regarder Laplace comme un précurseur des psychophysiciens. Dans le dit *Essai*, p. 1, il écrit : « Les questions les plus importantes de la vie ne sont pour la plupart que des problèmes de probabilité ! On peut même dire, à parler en rigueur, que presque toutes nos connaissances ne sont que probables ; et dans le petit nombre des choses que nous pouvons savoir avec certitude, dans les sciences mathématiques elles-mêmes, les principaux moyens de parvenir à la vérité, l'induction et l'analogie, sont fondés sur les probabilités ; en sorte que le système entier des connaissances humaines se rattache à cette théorie ».

(*) De 1815 à 1840, l'*Essai philosophique sur les probabilités* n'obtint pas moins de six éditions.

12. — Revenons au problème des partis. Il est exposé avec plus ou moins de détail dans presque tous les traités de calcul des probabilités. On a compliqué ce problème en fixant aux divers joueurs, $A_1, A_2, \ldots\ldots, A_n$, des nombres inégaux de points pour gagner : Le jeu se terminera soit quand A_1 aura obtenu s_1 points, soit quand A_2 aura s_2 points etc., soit enfin quand A_n aura s_n points. Et A_1 sera réputé gagnant et aura droit à toute la mise, s'il obtient s_1 points avant que A_2 n'en gagne s_2 etc., et avant aussi que A_n n'en gagne s_n ; en général, c'est le joueur A_λ qui aura droit à toute la mise, s'il gagne s_λ points avant qu'aucun des autres joueurs A_μ n'arrive aux s_μ points qui lui sont assignés ; les nombres s_1, $s_2, \ldots, s_\lambda, \ldots, s_\mu, \ldots, s_n$ étant des nombres entiers positifs quelconques.

On a généralisé encore le problème des partis en introduisant la possibilité de parties nulles. A chaque partie, il y a : probabilité mathématique p_1 pour que le joueur A_1 gagne 1 point, probabilité mathématique p_2 pour que A_2 gagne 1 point, et probabilité mathématique $r = 1 - p_1 - p_2$ pour que ces deux joueurs fassent « partie nulle » et soient considérés comme n'ayant gagné ni l'un ni l'autre. Le jeu prend fin soit quand A_1 a gagné s_1 points, soit quand A_2 a obtenu s_2 points ; et le premier qui atteint le nombre de points à lui assigné est réputé gagnant (17). Ce problème avec la complication des parties nulles se généralise facilement au cas d'un nombre quelconque n de joueurs.

Enfin, on a introduit une dernière complication : on suppose que la probabilité mathématique de gagner 1 point varie à chaque coup. Le joueur A_1 a la probabilité mathématique $p_{1\lambda}$ de gagner la $\lambda^{\text{ième}}$ partie, le joueur A_2 la probabilité mathématique $p_{2\lambda}$ etc., et le joueur A_n la chance $p_{n\lambda}$ de gagner la $\lambda^{\text{ième}}$ partie ($\lambda = 1, 2, 3, \ldots\ldots, n$). Et il y a probabilité mathématique

$$r_\lambda = 1 - p_{1\lambda} - p_{2\lambda} - \ldots\ldots - p_{n\lambda}$$

pour que les joueurs fassent partie nulle au $\lambda^{\text{ième}}$ coup. Comme ci-dessus, le jeu prend fin : soit quand A_1 est arrivé à s_1 points, soit

(17) Voir le grand traité de Mr Louis Bachelier, *Calcul des probabilités*, t. I, Paris, 1912, p. 42-43.

quand A_2 est arrivé à s_2 points ou A_3 à s_3 points etc. Le gagnant est toujours celui qui arrive le premier au nombre de points à lui assigné. Or, on interrompt le jeu, sans attendre la décision du sort, lorsque A_1 a déjà obtenu a_1 points ($a_1 < s_1$) et A_2 déjà a_2 points ($a_2 < s_2$) etc., enfin A_n déjà a_n points, où $a_n < s_n$. Comment partager alors équitablement la mise totale ?

Cette esquisse rapide de l'évolution du premier problème de la théorie des chances, le problème des partis, suffira ici pour illustrer *le courant mathématique*. Cet exemple montre l'ampleur prise par le calcul des probabilités au point de vue mathématique.

13. — Quant au côté *philosophique*, il ne s'est développé ni si vite, ni si magistralement. La cause en réside dans les difficultés réelles que présentent les définitions fondamentales, mais aussi dans le fait qu'au début, c'était la résolution de problèmes particuliers, bien précis, qui retenait l'attention. Il est intéressant de remarquer, au point de vue de la formulation des problèmes, qu'il s'agit presque toujours de déterminer *une valeur*, au sens pratique et commercial de ce terme. Toute chance apparaît comme une chance de gain, et déterminer cette chance n'est pas autre chose que déterminer le montant de la somme équivalente, de traduire la chance aléatoire en propriété effective et certaine. Pascal (18) dit une fois dans sa correspondance : « Le hasard est égal ». C'est peut-être la première fois qu'apparaît explicitement la notion de « cas également possibles », si importante pour la suite du développement et si souvent débattue.

La même formulation inspirée de la signification pratique se retrouve en général chez les premiers auteurs qui traitèrent la théorie des chances, Huygens et Jacques Bernoulli. On y lit fréquemment par exemple : l'expectative est dite valoir $\frac{1}{2}(a + b)$. Et il est curieux que beaucoup plus tard encore, Th. Bayes parte aussi de ce point de vue commercial (19).

(18) *Œuvres de Blaise Pascal*, La Haye, 1779, t. IV, p. 414.

(19) *An Essay towards solving a Problem in the Doctrine of Chances.* By the late Rev. Mr. Bayes. Communicated by Mr. Price. Philosoph. Transact. London, 1763 (publ. 1764), vol. 53, p. 370-418; ibidem, 1764 (publ. 1765), vol. 54, p. 296-325.

III. — *Le courant sociologique et la stocastique.*

14. — Bien des années après les premiers travaux, le calcul des probabilités recevait une nouvelle impulsion dans une tout autre direction. Elle venait cette fois de l'Angleterre et de la Hollande. C'est en effet de là que partit *le courant sociologique* mentionné au n° 1, courant orienté vers les applications pratiques telles que la statistique, les diverses branches de l'assurance etc. En Hollande, le problème suivant se posait : quel prix peut-on équitablement demander ou raisonnablement payer pour l'engagement de servir une rente viagère à une personne d'un âge donné x? Cette question revient manifestement à celle de la probabilité, pour un individu d'âge x, de vivre encore n années. Mais il s'agit de nouveau de déterminer une valeur au sens pratique et commercial. Les premiers essais de résoudre ce problème remontent aux auteurs suivants :

1° A John Graunt (1620-1674), marchand de la Cité de Londres. Dès 1592, à la suite d'une maladie contagieuse, les Anglais avaient pris l'habitude de dresser des relevés obituaires et depuis cette époque, avec quelques interruptions, les clercs des différentes paroisses publiaient régulièrement, le jeudi d'avant Noël, une liste générale des décès de l'année. C'est en se basant sur ces listes que John Graunt a essayé entre autres choses de dresser une sorte de table de mortalité, cherchant à déterminer la mortalité probable à chaque âge (20).

2° Au Grand-pensionnaire hollandais Jean de Witt (1625-1672), géomètre distingué autant qu'homme d'Etat tout à fait remarquable. Il adressa en 1671 aux Etats généraux un mémoire où il détermine, à l'aide des principes du calcul des probabilités, les valeurs des rentes viagères pour les comparer à celles des annuités certaines. C'est le premier ouvrage connu sur les rentes viagères ; il

(20) *Natural and political observations ..... made upon the bills of mortality.* London, 1662, ouvrage qui eut quatre éditions du vivant de son auteur ; une 5e fut publiée en 1676 par le chevalier William Petty (1623-1687) ; voir la note 24.

est instructif, même abstraction faite de son intérêt historique (21).

3° A Jean Hudde de Waveren (1633-1704), bourgmestre d'Amsterdam, fort bon géomètre également. Sa correspondance avec de Witt et Huygens nous révèle les intéressantes recherches sur la mortalité en Hollande, faites en collaboration avec Jean de Witt (22).

4° A Edmond Halley (1656-1742), célèbre astronome en même temps que mathématicien de rang. Il alla étudier à son tour, dans les relevés obituaires de Londres et de Breslau, les lois de la mortalité humaine et publia sur ce sujet un mémoire qu'on lit encore aujourd'hui avec admiration (23). La table de mortalité d'Edmond Halley a été tirée des observations faites à Breslau, pendant les cinq années 1687 à 1691, par l'ecclésiastique Caspar Neumann (1648-1715), à la demande de Halley. Ce dernier indique clairement pourquoi il a préféré, pour ce genre de recherches, la ville de Breslau à Londres et à Dublin (24).

15. — Malgré la nouveauté du sujet et la réputation des auteurs, ces ouvrages n'eurent pas le succès qu'ils méritaient. Le Calcul des probabilités, qui devait un jour faire un tel bruit dans le monde, y était entré au milieu de l'indifférence générale et, pour des raisons de divers ordres, elle se prolongea plus d'un demi-

(21) *Waerdye van Lyf-Renten naer proportie van Los-Renten*, s'Graven-Hage, 1671. Cet ouvrage, ayant été interdit après le massacre de Jean de Witt dans une émeute provoquée par les Orangistes, a de tout temps été fort rare. Il n'est devenu facilement accessible que depuis 1879, année où la Société hollandaise des sciences, à l'occasion de son centenaire, en fit paraître une traduction hollandaise.

(22) Voir *Œuvres complètes de* Christiaan Huygens, publiées par la Soc. holland. d. Sciences, t. VII, La Haye, 1897, p. 95.

(23) *An Estimate of the degrees of the mortality of mankind, drawn from curious Tables* ..... Philosoph. Transact., London 1693, vol. XVII, p. 596-610 ; ibidem, p. 654-656.

(24) Il est juste de dire que le chevalier William Petty, l'auteur des *Essays in political arithmetics*, London, 1699, dans différents écrits sur l'économie politique, s'était, de 1682 à 1687, occupé de semblables recherches. Mais Petty n'était pas assez mathématicien pour traiter la question fructueusement ; il y avait dans ses écrits plus d'imagination que de jugement ; cependant, Halley connut ces différents travaux et en profita.

siècle encore. Même les beaux travaux de Halley, enfouis dans la vaste et riche collection des mémoires de la Société royale de Londres, n'y devaient être découverts que par la postérité. Il fallut l'*Ars conjectandi* de Jacques Bernoulli, il fallut les immenses perspectives que ce livre ouvrait enfin à un calcul resserré soixante ans dans des limites fort étroites, pour frapper les esprits et donner à la découverte de Pascal et Fermat la popularité dont elle avait été si longtemps privée (25). Ce fut alors un élan vers l'étude des applications aux choses de l'économie politique et civile. La génération qui suivit Jacques Bernoulli se consacra à dépouiller des registres d'état-civil, à former des listes de naissances, de mariages, de décès, dans les différents pays et pour différentes époques, à compulser des documents d'économie politique etc.; bref, elle créa *la statistique*, science du dénombrement méthodique des faits susceptibles d'être comptés et qui a pour objet non seulement de recueillir, mais de coordonner ces faits, de manière à déterminer des rapports numériques sensiblement indépendants des anomalies du hasard. Il fallait pour ce travail plus de patience que de génie et plus de loisir que d'invention ; mais ce travail était indispensable, puisqu'il fournissait des matériaux où le Calcul des probabilités devait trouver un jour les éléments de ses plus importantes découvertes.

16. — L'application du calcul des probabilités à des domaines aussi disparates que les jeux de hasard, les rentes viagères, la mesure de la mortalité etc., firent apparaître l'importance et la généralité de la nouvelle théorie. Mais on se rendit compte aussi qu'il ne suffisait pas de compiler des faits et de multiplier des observations. Si incontestable qu'en fût la valeur intrinsèque, il

(25) *L'Ars conjectandi* ne fut publié à Bâle qu'en 1713, soit huit ans après la mort de l'auteur, par un de ses neveux, Nicolas Bernoulli (1687-1759). Cet ouvrage de 341 p. in-4° est en 4 parties ; la première (71 p.) est une reproduction du Traité sus-mentionné de Huygens (12), enrichie de commentaires et de nombreux résultats nouveaux. Traduction française : *L'Art de conjecturer*. Avec observations, éclaircissemens et additions. Traduit du latin de Jacques Bernoulli par L.-G.-F. Vastel, Caen, an X (1801).

fallait des méthodes permettant de les travailler et d'en tirer tous les renseignements utiles ; sinon, ces documents étaient condamnés à la stérilité. Alors commença une série de travaux remarquables. « Ils remplissent », dit Gouraud (4), « non seulement dans l'histoire de la théorie des hasards, mais dans celle de la pensée humaine, une page que le temps n'effacera jamais. Tout ce que le dix-huitième siècle avait de grands hommes descendit concurremment dans l'arène ». Le calcul des probabilités occupa de Moivre, Daniel Bernoulli, d'Alembert, Buffon, le grand Euler, Lagrange, Condorcet, Laplace, Poisson, Gauss, pour ne citer que les plus célèbres. *La stocastique* fut créée ou du moins préparée, science qui a pour objet l'application des théorèmes du calcul des probabilités aux nombres recueillis par la statistique, de manière à dénoter l'existence de causes régulières dont l'action se combine avec celle de causes fortuites.

Il était tout à fait dans l'esprit du dix-huitième siècle, et conforme au courant philosophique de l'époque, de mettre les plus grands espoirs sur le calcul des probabilités et de voir cette théorie rationaliser l'art de penser. En peu de temps, on arriva à une conception théorique qui est encore aujourd'hui généralement admise, du moins par les mathématiciens. Ce n'est plus la détermination d'une valeur commerciale qui est assignée comme but à la théorie ; celle-ci prend le caractère d'une méthode universelle et devient une logique générale de l'incertain.

IV. — *Le courant philosophique.*

17. — Jacques Bernoulli, dans l'*Ars conjectandi*, considère la connaissance comme une quantité ; la certitude est cette quantité entière, la probabilité en est une fraction. La certitude entière joue le rôle de l'unité ; on peut représenter la certitude par 1 et, dès lors, exprimer par toutes les parties fractionnaires de 1 les degrés infinis de probabilité par où l'esprit peut passer. Quand cette fraction devient infiniment petite, la limite vers laquelle elle tend, zéro, n'est plus pour l'intelligence que l'expression mathématique de l'impossible.

Cette métaphysique supposée rigoureuse, la vraisemblance est mise sous forme de fraction numérique et il en résulte la définition bien connue : *La probabilité mathématique d'un événement est égale au quotient obtenu en divisant le nombre des cas favorables à l'arrivée de cet événement par le nombre total des cas possibles, pourvu que les cas envisagés soient tous également possibles.*

Cette définition classique, qui pendant plus de deux siècles fut la base de toute la théorie des probabilités mathématiques, ne détermine pas la notion et ne dévoile nullement la nature du concept caché derrière le mot de probabilité ; elle ne fait que donner un précepte permettant d'exprimer numériquement, et sous des hypothèses très particulières, ce qu'il y a de quantitatif dans ce concept. La définition posée ne touche donc que l'un des côtés de la notion, celui qui forme l'objet de considérations mathématiques. C'est pour caractériser cette restriction que j'ajoute le qualificatif de « mathématique » au terme de probabilité.

De pareilles définitions, qui ne pénètrent pas jusqu'au centre d'une question, mais permettent néanmoins de saisir l'aspect quantitatif de certains phénomènes, ne sont pas rares dans les sciences. On en rencontre surtout en physique mathématique. Ainsi la définition habituelle de la vitesse : « Le chemin parcouru divisé par le temps employé à le parcourir », ou bien « La dérivée du chemin par rapport au temps », $ds : dt$, cela n'épuise pas la notion de vitesse, car lorsqu'on parle de la vitesse d'une réaction ou de la vitesse de refroidissement ou de la rapidité d'une action, il n'est nullement question de chemin parcouru ; on ne peut pas appliquer la définition classique $ds : dt$ à tous les phénomènes qui impliquent l'idée de vitesse. De même, la définition classique de la probabilité mathématique n'est pas applicable à tous les faits qui impliquent l'idée de probabilité. Une quantité d'épithètes en font foi : une probabilité peut être objective ou subjective, absolue ou relative, à priori ou à postériori, mathématique ou empirique, statistique, philosophique, psychologique, etc. Il a fallu deux siècles et demi pour démêler l'enchevêtrement de ces idées. L'un des résultats les plus certains, démontré irréfutablement, c'est l'existence de deux grandes catégories de probabilités, l'une comprenant les probabilités *anumériques*, l'autre les probabilités susceptibles de mesure

numérique. C'est parmi ces dernières seulement que se trouve la probabilité mathématique.

18. — Dans la définition classique de la probabilité mathématique rappelée ci-dessus, il faut attacher la plus grande importance à la notion de l'égale possibilité des cas. C'est de là qu'est parti *le courant philosophique* mentionné au n° 1, courant qui aboutit à la Logique et à la Théorie de la connaissance. Le point de vue de Jacques Bernoulli, adopté par tous les auteurs classiques de la théorie des chances, se base sur ce que j'appelle *le principe de la raison manquante* ou encore *le principe de l'absence de raison justifiante*. Il est formulé explicitement déjà par le grand géomètre bâlois Daniel Bernoulli (1700-1782), fils de Jean I Bernoulli, dans son « Esquisse d'une théorie nouvelle de la mesure des chances », parue en 1738 (26). On trouve ce point de vue marqué en termes encore plus nets dans l'article « Probabilité » de la fameuse *Encyclopédie* de Diderot et d'Alembert, où d'Alembert insiste sur le caractère subjectif de cette notion (27). C'est aussi le point de vue adopté par Laplace ; dans sa *Théorie analytique des probabilités* (v. n° 11), il donne cette brève explication : « Cas également possibles, c'est-à-dire tels que nous soyons également indécis sur leur existence » ; et il ajoute : « Leur juste appréciation est un des points les plus délicats de la théorie des hasards ».

D'après ce principe, des cas doivent être envisagés comme également possibles lorsque, dans l'état actuel de nos connaissances, nous ne pouvons discerner aucune raison d'envisager l'un comme plus probable qu'un quelconque des autres ; à part la disjonction des cas, il y a ignorance absolue. On a qualifié ce principe, non

(26) *Specimen theoriae novae de mensura sortis.* Comment. Acad. scient. imper. Petropolitanae, vol. V (1730-1731), publ. 1738, p. 175-192. Trad. allem. avec notes par A. Pringsheim, Leipzig, 1896. C'est dans ce mémoire que Daniel Bernoulli, à propos du problème appelé « problème de Saint-Pétersbourg », développe sa fameuse théorie de « l'espérance morale », par opposition à celle de « l'espérance mathématique ».

(27) *Encyclopédie, ou Dictionnaire raisonné des sciences, des arts et des métiers,* publié par M. Diderot et M. d'Alembert. Paris, 1751-1769.

sans raison, de « principe de la distribution uniforme de l'ignorance ».

Envisageons comme exemple une urne ; supposons qu'on sache qu'elle contient seulement des boules blanches et des boules noires, mais qu'on ne sache rien de plus de son contenu. En en tirant une boule au hasard, on a donc deux possibilités : sortir boule blanche ou sortir boule noire. D'après le principe de l'absence de raison justifiante, ces deux possibilités constituent des cas également vraisemblables ou *équivraisemblables*, puisque les renseignements que l'on possède ne permettent pas de supposer qu'il se trouve dans l'urne plus de boules d'une des deux couleurs que de l'autre. Cette doctrine a toujours trouvé de très nombreux adeptes jusqu'à nos jours ; un auteur moderne va même jusqu'à poser comme définition : Le calcul des probabilités est une méthodisation de l'ignorance (28) (v. n° 125, remarque, et n° 143).

19. — Diamétralement opposé à ce point de vue est celui du *principe de la raison péremptoire*, formulé pour la première fois très nettement par Jean de Kries (29). Selon les philosophes de cette école, on ne peut parler de cas également possibles ou *équipossibles* que si des faits ou certaines considérations nous obligent à les considérer comme également enclins à se produire. S'il s'agit de faits physiques, l'équipossibilité doit reposer sur des connaissances objectives qui excluent tout arbitraire.

On ne peut s'empêcher de donner raison à ces philosophes-là.

En effet, par quelle sorcellerie l'Analyse mathématique voudrait-elle tirer quelque chose du néant ? Elle ne peut que déduire de prémisses données ce qu'elles contiennent déjà implicitement ; et lorsqu'il y a ignorance absolue sur un domaine D, on ne voit pas comment le maniement de symboles algébriques ou de formules du calcul intégral, si élégantes soient-elles, permettrait de tirer des connaissances positives sur ce domaine D.

(28) Samuel Lourié, *Die Prinzipien der Wahrscheinlichkeitsrechnung*, eine logische Untersuchung des disjunktiven Urteils. Tübingen, 1910.

(29) Johannes von Kries, *Die Prinzipien der Wahrscheinlichkeitsrechnung*. Eine logische Untersuchung. Freiburg i. Br., 1886.

Reprenons l'exemple de l'urne. On ne pourra considérer les cas de sortie de boule blanche ou sortie de boule noire comme également enclins à se produire, et les poser comme équipossibles, que si l'on *sait* qu'il se trouve dans l'urne autant de boules blanches que de noires ; car dans ces conditions, je suis *obligé* de considérer les deux cas comme équipossibles.

Du point de vue purement logique, les deux principes sont certainement admissibles. Leur importance apparaît surtout quand il s'agit d'appliquer la théorie des chances basée sur la définition classique. Si l'on adopte le premier principe, celui de l'absence de raison justificative, on verra qu'il n'existe guère de domaine d'investigation où il soit impossible de statuer des cas « également vraisemblables », en se basant sur un manque absolu de connaissances. Si l'on se place au contraire sur le terrain du principe de la raison péremptoire, on sera amené à restreindre considérablement le domaine d'applicabilité de la théorie des chances.

20. — La phase moderne du courant philosophique commence avec les travaux d'Antoine-Augustin Cournot (1801-1877), grand penseur français, à la fois mathématicien, économiste, philosophe et historien (30). Né à une époque où le mouvement en faveur du calcul des probabilités ne se limitait pas à la France, Cournot vécut au milieu d'une génération imprégnée de probabilisme. Il est l'auteur de plusieurs ouvrages tout à fait remarquables par la clarté limpide du style et l'originalité des idées (31). Contrairement à Laplace qui niait le hasard au nom du déterminisme scientifique, Cournot proclame la parfaite conciliabilité de ces deux notions en apparence opposées. Pour Laplace comme pour les Encyclopédistes, le hasard n'est qu'un « vain son », un mot servant à déguiser ou à dissimuler l'ignorance où nous sommes des véritables causes,

(30) Voir dans la Revue de métaphysique et de morale, t. XIII, p. 293-543, Paris, 1905, un numéro spécialement consacré à Cournot, notamment l'article de Mr F Mentré, *Les racines historiques du probabilisme.*

(31) A.-A. Cournot, *Exposition de la théorie des chances et des probabilités*, Paris, 1843, VIII + 448 p. in-8° + 1 planche de fig. Ce livre se rapporte plus particulièrement à notre sujet.

car dans l'Univers, tout est rigoureusement déterminé. « Nous devons envisager l'état présent de l'Univers comme l'effet de son état antérieur et comme la cause de celui qui va suivre », dit Laplace. Depuis Spinoza (1632-1677), la majorité des philosophes accepte ce point de vue, et Bossuet (32) écrit : « Ne parlons donc plus de hasard ni de Fortune, ou parlons-en seulement comme d'un nom dont nous couvrons notre ignorance. Ce qui est hasard à l'égard de nos conseils incertains est un dessein concerté dans un Conseil plus haut, c'est-à-dire dans ce Conseil éternel qui renferme toutes les causes et tous les effets dans un même ordre ».

Chez Voltaire (33) et les Encyclopédistes, la tournure théologique disparaît : « Ce que nous appellons hasard n'est et ne peut être que la cause ignorée d'un effet connu ». Parmi les très nombreux auteurs adoptant ce point de vue, il convient de citer, à cause de la grande influence de ses écrits, l'économiste belge Quételet (34). « Le mot hasard », dit-il, « sert officieusement à voiler notre ignorance ». La réalité objective du hasard est donc niée ; on n'admet qu'une ignorance, purement subjective, de la nature des choses ou des lois qui les régissent ; ou bien, quand ces lois sont connues, une ignorance due à la complication excessive des conditions dont les événements dépendent. Selon tous ces penseurs, la Science tend à éliminer de plus en plus les faits fortuits et, à la limite, c'est-à-dire pour un Esprit universel et omniscient, il n'y aurait plus ni hasard ni probabilité ; il ne pourrait y avoir que certitude.

Pour Cournot, au contraire, le hasard et par conséquent la probabilité mathématique ont une réalité objective et indépendante de l'état de nos connaissances. Ils subsisteraient même pour une Intelligence supérieure capable d'embrasser l'univers tout entier. Cournot a su discerner l'*élément logique* dans la notion de hasard. Il a si bien vu l'importance primordiale de cet élément qu'il est parti de là pour définir le hasard. Cournot n'utilise pas encore le terme de « hasard logique », mais il en a déjà la notion, fonde-

(32) Jacques-Bénigne Bossuet (1627-1704), *Discours sur l'histoire universelle*, III, 8, 1681.

(33) Voir dans le *Dictionnaire philosophique* de Voltaire, l'article Athéisme.

(34) Voir la note (88) du n° 157 et la note (100) du n° 186.

ment de la définition à laquelle nous aboutirons pour la probabilité mathématique (v. n° 135, définition 8, n°s 136 et suivants)

21. — Pour asseoir sa théorie philosophique des probabilités sur une base inattaquable, il manquait encore à Cournot deux choses : les notions si précises de mesure que seule la théorie moderne des ensembles permet de définir en toute rigueur, et la méthode axiomatique menée à un très haut degré de perfection par les travaux récents sur l'arithmétisation des mathématiques.

Les premiers essais d'appliquer la méthode axiomatique au calcul des probabilités, pour en faire une science ne donnant plus prise aux objections qu'on ne cessait de soulever contre la valeur philosophique et pratique de ses résultats, remontent aux géomètres suivants : 1° à Mr Bohlmann, auteur de l'article sur les mathématiques d'assurances dans l'*Encyclopédie des sciences mathématiques pures et appliquées* (35) ; 2° à Mr Broggi qui a fait sa thèse de doctorat sur ce sujet (36) ; 3° à Mr Emile Borel qui a consacré à cette question une partie d'un mémoire de 1909 (37). Enfin 4° à Mr de Mises qui, par une innovation radicale, semble avoir réussi à doter le calcul des probabilités d'une base logiquement irréprochable (38).

Pour réussir dans son entreprise difficile, Mr de Mises abandonne hardiment la définition classique de la probabilité mathématique. La définition nouvelle qu'il adopte s'inspire des idées théoriques de Cournot aussi bien que des résultats de la stocastique dont Fechner marqua l'avènement par sa théorie des ensembles statistiques, théorie qu'il baptisa de *Kollektivmasslehre* (39), développée

(35) Tome I, volume 4, fasc. 4, p. 496-501, Paris et Leipzig, 1911, *Technique de l'assurance sur la vie*. Exposé, d'après l'article allemand de G. Bohlmann, par H. Poterin du Motel.

(36) Hugo Broggi, *Die Axiome der Wahrscheinlichkeitsrechnung*, Inauguraldiss., Göttingen, 1907.

(37) Emile Borel dans *Rendiconti* del Circolo matematico di Palermo, t. 27, 1909, p. 247-271.

(38) Richard Edler von Mises, *Grundlagen der Wahrscheinlichkeitsrechnung*, Math. Zeitschrift, t. 5, Berlin, 1919, p. 52-99.

(39) Gustav Theodor Fechner (1801-1887) *Kollektivmasslehre*. Œuvre posthume éditée par G. F. Lipps, Leipzig, 1887.

en Allemagne surtout par Mr Bruns (40). Le calcul des probabilités n'est plus ni une méthode de déterminer une valeur commerciale, comme dans le premier stade (v. nº 13), ni une logique générale de l'incertain, plus métaphysique que mathématique, comme dans le deuxième stade du développement (v. nº 17) ; c'est une discipline mathématique au même titre que par exemple la Mécanique rationnelle ou la Géométrie, discipline qui s'applique à la réalité physique avec un degré plus ou moins grand d'exactitude, comme toute autre théorie mathématique.

Cette manière moderne d'asseoir le calcul des probabilités constitue en quelque sorte une inversion de la théorie classique des chances ; en effet, en partant de ce nouveau point de vue, on arrive à *démontrer*, sous certaines hypothèses, la formule $p = f : s$ posée comme définition au nº 23. La définition classique devient ainsi un théorème (v. nº 197) ; et cela élimine du même coup les grandes difficultés philosophiques introduites dans la théorie habituelle par la notion de l'égale possibilité des cas. Avant Mr de Mises, Mr Helm (40*a*) avait déjà eu l'idée de renverser ou plutôt « inverser » la définition de la probabilité mathématique, c'est-à-dire de la ramener à la mesure et à l'étude de certains ensembles appelés collectifs, au lieu de la baser sur des cas « équipossibles ». Mais c'est Mr de Mises qui le premier a mené cette idée à chef au point de vue mathématique. Nous aurons à revenir sur les diverses phases de l'évolution des concepts fondamentaux ainsi que sur les trois courants mentionnés au nº 1 et dont l'action combinée a conquis au calcul des probabilités une place de plus en plus centrale dans presque toutes les sciences.

(40) Heinrich Bruns, *Wahrscheinlichkeitsrechnung und Kollektivmasslehre*, Leipzig et Berlin, 1906.

(40*a*) R. Helm, *Die Wahrscheinlichkeitslehre als Theorie der Kollektivbegriffe*, dans Annalen der Naturphilosophie, t. I, 1902, p. 364-381.

CHAPITRE II

LES ÉLÉMENTS DU CALCUL DES PROBABILITÉS

I. — *La définition classique de la probabilité mathématique.*

22. — Pour parvenir à la définition habituelle de la probabilité mathématique, nous allons raisonner sur un exemple concret. Un objet de valeur est mis en loterie ; elle se compose de 1 000 numéros vendus aux joueurs. On mélange dans une urne 1 000 billets pareils entre eux, numérotés de 1 à 1 000 ; puis on en tire un au hasard. Le lot échoit au joueur dont le numéro est précisément celui du billet sorti. Dans ce cas, il est naturel de raisonner comme suit.

a) Plus un joueur possède de numéros, plus il a de chances de gagner. S'il a acheté 2, 3, 4, ..., n fois plus de numéros qu'un autre, il a aussi 2, 3, 4, ..., n fois plus de chances de gain. Donc, la probabilité de gagner doit être posée proportionnelle au nombre f des numéros achetés. La chance de gagner s'exprimera dès lors par $k'.f$, où k' désigne un facteur de proportionnalité.

b) Supposons que trois objets de valeur aient été mis en loterie de cette manière, chacun donnant lieu à une loterie distincte des autres. Pour la première, on a fait 1 000 numéros, pour la seconde seulement 100 et pour la troisième 4 000 numéros ; puis, tous les numéros vendus, on procède aux tirages.

Envisageons un joueur qui achète un numéro dans chacune de ces trois loteries. Tout le monde s'accordera à dire que sa chance de gagner est 10 fois plus grande à la seconde loterie qu'à la première, puisqu'il possède 1 % de tous les numéros de la seconde et seulement $\frac{1}{10}$ % des numéros de la première et que tous les

numéros ont *a priori* la même chance d'être tirés. Pour la même raison, sa chance de gagner est 4 fois plus petite à la troisième loterie qu'à la première. D'une manière générale, si le nombre de tous les numéros devient 2, 3, 4, ..., n fois plus grand, la chance de gagner pour le possesseur d'un seul billet devient 2, 3, 4, ..., n fois plus petite. Donc, la probabilité de gagner doit être posée inversement proportionnelle au nombre total des numéros ; elle s'exprimera dès lors par $k'' : s$, où k'' représente de nouveau un facteur de proportionnalité.

c) En combinant ceci avec le résultat obtenu sous **a**), on voit que, si l'on veut une science conforme au bon sens commun, la probabilité de gagner doit être à la fois directement proportionnelle au nombre des billets achetés et inversement proportionnelle au nombre total des numéros. Dès lors, on doit la poser $= k.f : s$.

d) Envisageons une loterie où tous les billets soient gagnants. Alors $f = s$ et la probabilité de gagner devient $= k$. On voit par là que le facteur de proportionnalité, k, représente la certitude de gagner. On est libre de lui attribuer une valeur arbitraire. Pour des raisons de simplicité, on pose $k = 1$. La certitude est alors représentée par le nombre 1 et la probabilité mathématique de gagner, p, prend la forme

$$p = \frac{f}{s} . \qquad \text{(II 1)}$$

23. — Si l'on désigne par E l'événement de gagner l'objet mis en loterie, on voit qu'il y a s « cas possibles », puisque l'urne contient s billets et qu'on en tire seulement 1. Les f numéros achetés par le joueur sont autant de « cas favorables » à l'événement E En effet, le joueur gagnera dans f cas différents, ceux où le numéro sorti de l'urne coïncidera avec l'un ou l'autre des numéros qu'il possède.

D'une manière générale, soit E un événement aléatoire quelconque et $p(E)$ sa probabilité mathématique. Supposons que E se réalise dans f cas, à l'exclusion des $s - f$ autres, s étant le nombre de tous les cas, équipossibles par hypothèse. La formule

$$p(E) = p = f : s \qquad \text{(II 1)}$$

est l'expression algébrique de la définition formulée ci-dessus au n° 17 :

Définition 1. — La probabilité mathématique d'un événement aléatoire E *est le rapport du nombre des cas favorables à l'arrivée de* E, *au nombre total des cas possibles, pourvu qu'ils soient tous équipossibles.*

Problème 1. — En lançant au hasard un dé à jouer, quelle est la probabilité mathématique d'amener le point 3 par exemple? *Solution.* Il y a 6 cas possibles, puisque le dé a 6 faces, numérotées de 1 à 6; donc $s = 6$. Une seule porte le point 3; donc $f = 1$, et la probabilité mathématique cherchée est

$$p(3) = f : s = 1/6.$$

Remarque. — Il est important de noter que cette solution présuppose qu'on joue avec un dé idéal, c'est-à-dire géométriquement régulier et physiquement homogène, de sorte que son centre de gravité coïncide exactement avec son centre de figure. Sinon, les 6 cas ne seraient pas équipossibles et la définition classique deviendrait inapplicable. Dans ce qui suit, nous supposons toujours, sans en faire mention chaque fois explicitement, que les dés utilisés sont « bons », c'est-à-dire remplissent ces conditions de régularité et d'homogénéité parfaites (v. n° 27 et n° 189, exemple et remarque 1).

Problème 2. — Un gain échoit à un joueur s'il amène le sonnez en lançant au hasard 2 dés (v. n° 4, note p. 6). Quelle est sa probabilité mathématique de gagner? *Solution.* Le point 1 du premier dé peut se combiner avec l'un quelconque des 6 points du second dé, ce qui fait 6 possibilités; de même le point 2 du premier dé, de même encore le point 3 et ainsi de suite. Cela fait en tout $6.6 = 36$ cas possibles; donc $s = 36$. De ces 36 arrangements, un seul est favorable à l'événement E qui consiste ici à amener le sonnez : il faut que le premier dé marque 6 et le second également 6. Donc $f = 1$ et par conséquent

$$p(E) = f : s = 1 : 36.$$

Définition 2. — Lorsque l'événement étudié E ne se réalise pas, on dit qu'il y a réalisation de l'*événement contraire*. Désignons-le par F. Il y a donc $(s - f)$ cas favorables à F ; sa probabilité mathématique est dès lors

$$p(F) \equiv q = \frac{s-f}{s} = 1 - \frac{f}{s} = 1 - p. \qquad \text{(II 2)}$$

On appelle q la *probabilité complémentaire* de p. Il en découle ce

Théorème. La somme des probabilités mathématiques d'un événement aléatoire E *et de l'événement contraire* F *est égale à* 1. En formule :

$$p + q = 1. \qquad \text{(II 2')}$$

Exemple. Quelle est, pour le joueur du problème 2, la probabilité mathématique de perdre ? Elle est égale à la probabilité mathématique de ne pas amener le sonnez. C'est l'événement contraire du précédent ; sa probabilité mathématique est dès lors

$$p(F) = 1 - 1/36 = 35/36.$$

Problème 3. — Déterminer la probabilité mathématique de sortir boule blanche d'une Urne $U(a, b ; s)$ contenant a boules blanches, b noires, au total $s = a + b$ boules. *Solution.* Il y a s cas également possibles, voir n° 33, définition 8. L'événement E consistant ici à sortir une boule blanche, a cas lui sont favorables ; donc, la probabilité mathématique cherchée est $p(E) = \frac{a}{s}$.

24. — La définition de la probabilité mathématique posée au n° 23 donne lieu à plusieurs observations dont nous allons examiner les principales.

Première observation. On a souvent critiqué la définition classique parce qu'elle exige de savoir reconnaître des événements « également possibles ». Or dans ce cas, également possible signifie également probable ; la probabilité mathématique est ainsi définie par « l'égale probabilité », en d'autres termes : la probabilité est

définie par la probabilité, c'est un cercle vicieux. Henri Poincaré en particulier pensait qu'il n'est guère possible de donner une définition satisfaisante de la probabilité.

Réfutation. 1) Pour constater l'égalité de deux grandeurs, il n'est nullement nécessaire de les mesurer ni même de définir une unité de mesure. 2) Même si l'on admettait le cercle vicieux, la définition classique aurait l'avantage d'économiser de la réflexion ; car il est plus facile, en pratique, de décider si deux événements sont équipossibles que de définir le degré de probabilité d'un phénomène ; de même qu'il est plus facile, en pratique, de décider si deux surfaces sont également éclairées que de définir abstraitement l'intensité de l'éclairement ; de même qu'il est plus simple, en pratique, de décider si deux corps ont une égale température que de définir *in abstracto* la température d'un corps ou la température en soi. 3) En tout cas, la définition classique de la probabilité mathématique ramène une notion à une autre plus simple, ce que fait en dernière instance toute définition.

25. — Deuxième observation. La définition classique ne résulte nullement d'une nécessité logique, mais plutôt d'une manière naturelle de raisonner. C'est une définition conventionnelle et partant plus ou moins arbitraire, comme le sont les définitions en général. Au point de vue logique, on aurait tout aussi bien pu admettre, par exemple, que la probabilité de gagner est proportionnelle au carré du nombre des cas favorables et dire, par définition, qu'elle devient 4, 9, 16, 25, ..., n^2 fois plus grande, lorsque le joueur achète 2, 3, 4, 5, ..., n fois plus de numéros, et qu'elle devient n^2 fois plus petite, si le nombre total des numéros, s, devient n fois plus grand. On serait alors amené à définir :

$$p(E) = f^2 : s^2.$$

Au lieu de la seconde puissance, on pourrait prendre la troisième ou la quatrième ou un exposant n quelconque et poser $p(E) = f^n : s^n$ ou encore plus généralement $p(E) = f^n : s^m$. On s'est arrêté à la formule (II 1), qui correspond au cas $n = m = 1$, pour plusieurs raisons : 1] C'est de toutes les formules la plus simple, la plus naturelle et qui conduit aux calculs les plus faciles. 2] Elle est con-

forme à l'idée de valeur commerciale qui, dans le premier stade du développement, s'attachait à la notion de probabilité mathématique (v. n° 13). 3] Les résultats qu'elle fournit sont confirmés par l'expérience. 4] Nous verrons plus bas une raison qui, en donnant l'explication de cette concordance expérimentale, oblige à adopter la définition (II 1) (v. n° 136). Rendons-la pour l'instant plus plausible encore en développant, d'après Cournot (31), le point mentionné sous 2].

Quand un objet commerçable a été mis en loterie, chacun des billets qui représente un droit éventuel à cet objet peut à son tour être mis dans le commerce, et sa valeur vénale sera celle de la chance ou du droit éventuel dont ce billet est le signe. Il n'y a aucune raison pour attribuer à l'un des billets plus de valeur qu'à un autre ; par conséquent, deux personnes possédant l'une a, l'autre b billets, posséderont des valeurs qui seront entre elles dans le rapport de a à b. Et si l'on imaginait un marché où règne la libre concurrence, il s'établirait un cours pour les billets de loterie comme pour toute chose qui s'achète et se vend habituellement, et le prix de chaque billet ou chance serait inversement proportionnel au nombre total des billets. En effet : si le cours assignait à chaque billet une valeur moins élevée, le possesseur de l'objet mis en loterie n'en retrouverait pas le prix courant et par conséquent ne choisirait pas cette voie pour s'en défaire. Si au contraire le cours assignait à chaque billet une valeur plus élevée, les spéculateurs trouveraient avantage à acheter directement l'objet mis en loterie ou des objets du même genre, pour en distribuer ensuite la valeur sur des billets ou chances négociables ; et par leur concurrence, ils feraient baisser le cours jusqu'à ce qu'il se trouvât ramené au taux qu'il avait momentanément dépassé.

26. — Troisième observation. Il convient d'insister sur un point : pour l'établissement de cas équipossibles, la disjonction des cas doit être complète. Un exemple d'intérêt historique le montrera.

Problème 4. — Quelle est la probabilité mathématique, en lançant deux dés à jouer, d'amener une somme de points donnée d'avance ? *Solution.* Le nombre des cas possibles est $s = 36$

(v. n° 23, problème 2). Le tableau suivant montre la manière dont les sommes de points se répartissent sur les nombres 2, 3, 4, ..., 11, 12 ; on désigne par I le premier et par II le deuxième dé.

$n=2$	$n=3$	$n=4$	$n=5$	$n=6$	$n=7$	$n=8$	$n=9$	$n=10$	$n=11$	$n=12$
I II	I II	I II	I II	I II	I II	I II	I II	I II	I II	I II
$1+1$	$1+2$	$1+3$	$1+4$	$1+5$	$1+6$	$2+6$	$3+6$	$4+6$	$5+6$	$6+6$
	$2+1$	$2+2$	$2+3$	$2+4$	$2+5$	$3+5$	$4+5$	$5+5$	$6+5$	
		$3+1$	$3+2$	$3+3$	$3+4$	$4+4$	$5+4$	$6+4$		
			$4+1$	$4+2$	$4+3$	$5+3$	$6+3$			
				$5+1$	$5+2$	$6+2$				
					$6+1$					
$f=1$	$f=2$	$f=3$	$f=4$	$f=5$	$f=6$	$f=5$	$f=4$	$f=3$	$f=2$	$f=1$

Désignons par $p(n)$ la probabilité mathématique d'amener n points ; on a, en appliquant la définition (II 1),

$$p(2)=p(12)=\frac{1}{36}\,;\quad p(3)=p(11)=\frac{2}{36}=\frac{1}{18}\,;$$
$$p(4)=p(10)=\frac{3}{36}\,;\quad p(5)=p(9)=\frac{4}{36}=\frac{1}{9}\,;$$
$$p(6)=p(8)=\frac{5}{36}\,;\quad p(7)=\frac{6}{36}=\frac{1}{6}\,.$$

Leibniz, dans son premier travail mathématique (41), dit que la somme 11 ne peut se produire que d'une manière, l'un des dés

(41) G.W. Leibniz, *Dissertatio de arte combinatoria*, 1666. Voir Leibnizens mathemat. Schriften, herausgeg. von C. I. Gerhardt, IIe sect. vol. 1, Halle, 1858. Cette dissertation est plutôt philosophique et métaphysique que mathématique et bien inférieure au *Triangle arithmétique* de Pascal cité dans la note 11.

devant marquer 5 et l'autre 6 ; et que la somme 7 peut se produire de trois manières : l'un des dés peut marquer 1 ou 2 ou 3, l'autre doit alors amener 6 ou 5 ou 4. L'erreur est manifeste et provient d'une disjonction incomplète des cas possibles ; le tableau ci-dessus démontre que les cas envisagés par Leibniz comme équipossibles ne le sont pas en réalité (v. n° 9 et n° 136).

Problème 5. — Deux joueurs, A et B, également habiles, lancent des boules vers le même but. Le gagnant est celui dont la boule arrive le plus près du but ; A a deux boules à lancer, B une seule. Quelle est pour A la probabilité mathématique de gagner ?
Solution. Désignons par a_1, a_2, les points d'impact des deux boules lancées par A, et soit b celui de la boule lancée par B. En considérant les distances de ces points au but, on a les possibilités suivantes, où la dernière lettre désigne le point le plus rapproché du but et la première lettre le point qui en est le plus éloigné :

$$a_1a_2b,\ a_1ba_2\ ;\quad a_2a_1b,\ a_2ba_1\ ;\quad ba_1a_2,\ ba_2a_1.$$

Ces six cas sont par hypothèse équipossibles, puisque les joueurs sont réputés également habiles ; donc $s = 6$. On voit que 4 cas font gagner A, le second et les trois derniers ; donc $f = 4$. La probabilité demandée a dès lors pour mesure $p = 4 : 6 = 2 : 3$.

Remarque. — On pourrait aussi raisonner comme suit. Il y a 4 possibilités à envisager, $s = 4$; ce sont :

1) a_1 plus loin que b et a_2 plus loin que b ;
2) a_1 » loin » b et a_2 » près » b ;
3) a_1 » près » b et a_2 » loin » b ;
4) a_1 » près » b et a_2 » près » b.

Les trois dernières font gagner A ; donc $f = 3$ et par conséquent la probabilité mathématique cherchée est $p = 3 : 4 = 0{,}75$. Mais ce raisonnement serait faux, parce que les 4 cas ne sont pas également enclins à se réaliser. En effet, 1) et 4) représentent chacun deux cas ; dans 1), la succession des boules peut être a_1a_2b ou a_2a_1b, et dans 4) on peut avoir ba_1a_2 ou ba_2a_1. En tenant compte de cela, on retrouve $p = 2/3$.

27. — Sans la restriction que tous les cas doivent être *équipossibles*, la définition classique de la probabilité mathématique n'aurait guère de sens, car on pourrait alors raisonner comme d'Alembert à propos de n'importe quel événement aléatoire (42) : Il peut se faire que l'événement arrive, il se peut aussi que l'événement n'arrive pas ; ce sont deux cas possibles, $s = 2$; un seul est favorable à la réalisation de l'événement, donc $f = 1$. La probabilité mathématique serait ainsi toujours $= 1/2$. Bien que d'Alembert ait refusé de passer outre, son erreur est grossière ; avant de compter les chances, il faut s'assurer qu'elles sont équivalentes (Cf. n^{os} 18 et 19).

Si l'on base le calcul des probabilités sur la définition classique, c'est dans la formation des cas équipossibles, c'est-à-dire dans la décomposition de la matière du problème en cas simples, équivalents et également enclins à se réaliser, que consiste l'art propre de la Théorie des chances ; le reste n'est qu'une question d'arithmétique ou d'analyse mathématique. Avec la définition moderne de la probabilité mathématique, basée sur la notion de fréquence et de « syllepte », il n'est pas nécessaire de s'occuper de cette subtilité-là, et c'est un grand avantage. De plus, la nouvelle définition s'applique sans autre à une foule de cas où la classique tombe en défaut précisément à cause de l'impossibilité de constituer directement des « cas équipossibles », par exemple lorsqu'il s'agit d'un dé pipé et irrégulier ou de phénomènes démographiques (v. chap. VIII, en particulier n^{o} 189, remarque 1).

28. — Pour ramener au schéma de l'urne certaines applications du calcul des probabilités à la statistique, j'adopterai la convention suivante. A chaque habitant d'un pays ou de la contrée à étudier, j'assigne une urne, *son urne de vie*, contenant des boules

(42) Jean Lerond d'Alembert (1717-1783), *Réflexions sur le calcul des probabilités*. Opuscules math., Paris, t. II (1761), p. 1-25 ; t. IV (1768), p. 73-79, 283-310 ; t. VII (1780), p. 39-60, 384-386. *Doutes et questions sur le calcul des probabilités*, Mélanges de litt., d'hist. et de philosophie, Amsterdam, t. V (1773), p. 273-304. Comparer les articles *Croix ou pile*, et *Gageure*, dans l'Encyclopédie citée dans la note 27.

blanches et noires dans une proportion déterminée, connue ou inconnue, qui dépend entre autres : de l'âge de l'individu, de son sexe, du pays et de la contrée qu'il habite, de la profession qu'il exerce et de son genre de vie. Le jour de l'An, chacun plonge la main dans son urne de vie et en tire au hasard une boule. Si elle est blanche, c'est le signe qu'il vivra encore à la fin de l'année ; si la boule sortie est noire, il mourra dans le courant de l'année. Voilà l'hypothèse implicite sur laquelle sont fondées de très nombreuses applications du calcul des probabilités à la statistique démographique. Quant aux *notations*, nous désignerons toujours l'âge, x par exemple, par un indice au bas et à droite du symbole principal l ou d ou p ou q ou a etc., l désignant un nombre de vivants, d un nombre de décès, p une probabilité mathématique de vie, q une probabilité de décès etc. (43).

Ainsi, l'urne de vie d'une personne d'âge x contiendra par hypothèse l_x boules dont d_x noires et par conséquent $(l_x - d_x)$ blanches. Une année plus tard, l'urne contiendra l_{x+1} boules dont d_{x+1} noires, et ainsi de suite.

Problème 6. — Quelle est la probabilité mathématique pour une personne d'âge x de décéder au cours d'une année, les hypothèses ci-dessus étant admises ? *Solution.* En vertu des dites hypothèses, la question posée revient à cette autre : quelle est la probabilité mathématique de tirer une boule noire d'une urne $U(l_x - d_x,\ d_x;\ l_x)$ contenant d_x noires et $(l_x - d_x)$ blanches ? Par définition (v. n° 23), la probabilité cherchée, que l'on désigne par q_x, est (43) :

$$q_x = d_x : l_x \qquad \text{(II 3)}$$

L'événement contraire, extraction d'une boule blanche de l'urne de vie, a comme probabilité mathématique

$$p_x = \frac{l_x - d_x}{l_x} = 1 - \frac{d_x}{l_x} = 1 - q_x \qquad \text{(II 4)}$$

(43) Voir L.-G. Du Pasquier, *Introduction à la science actuarielle*, Neuchâtel et Paris, 1918.

Définitions 3 et 4. — On appelle ce p_x le *taux annuel de vitalité* et ce q_x le *taux annuel de mortalité* correspondant à l'âge x. Conformément au théorème du n° 23, on a

$$p_x + q_x = 1 \qquad \text{(II 5)}$$

Ajoutons que depuis Euler, les actuaires ont développé des méthodes qui permettent de déterminer assez exactement les nombres d_x, l_x, p_x, q_x ; ces grandeurs interviennent dans une foule de problèmes importants (v. n° 14 et n° 16).

29. — Quatrième observation. Pour pouvoir être interprétée comme probabilité mathématique au sens de la définition classique, une fraction comprise entre 0 et 1 doit satisfaire à une condition formelle souvent négligée dans les applications : *L'ensemble des f cas dénombrés au numérateur doit faire partie de l'ensemble des s cas dénombrés au dénominateur* ; en d'autres termes, au numérateur ne doit figurer aucun cas ne se trouvant pas aussi parmi ceux comptés au dénominateur. Dans les applications à la statistique, il ne faut pas oublier cette condition préalable. On va le voir sur un exemple. Supposons qu'il s'agisse de déterminer le taux annuel de mortalité des nouveau-nés, q_0, de la population suisse. Appliquons une méthode due à Euler, pour dresser une table de mortalité.

1° Un recensement opéré au commencement de l'année, mettons le premier janvier 1900, fait connaître l_0, nombre des enfants suisses âgés de moins de douze mois à ce moment-là.

2° A l'aide des registres d'état civil, on peut connaître d_0, nombre des décès arrivés en Suisse, au cours de l'année 1900, parmi les enfants de moins d'un an.

3° On pourrait en conclure que le taux de mortalité cherché se détermine par simple division, en vertu de la formule (II 3) qui donne

$$q_0 = d_0 : l_0 \qquad \text{(II 6)}$$

Mais cette conclusion serait erronée. En effet, l'ensemble des d_0 décès comptés au numérateur ne provient pas de l'ensemble des l_0 nouveau-nés comptés au dénominateur, et cela pour trois raisons :

α) A cause de l'*immigration*. Parmi les d_0 décès figurent ceux provenant des enfants de familles immigrées en Suisse au cours de l'année 1900 ; n'ayant pas été atteints par le recensement du premier janvier 1900, ces enfants ne sont pas compris dans le nombre l_0.

β) A cause de la *natalité*. Parmi les d_0 décès figurent ceux d'enfants nés en Suisse après le premier janvier 1900 et morts en Suisse avant le 31 décembre 1900 ; ceux-là ne figurent pas non plus dans le nombre l_0.

γ) A cause de l'*émigration*. Un certain nombre de bébés, recensés le premier janvier 1900 et morts avant d'atteindre un an, ne sont tout de même pas comptés parmi les d_0 cas, leur décès s'étant produit en dehors du territoire suisse.

Pour ces trois raisons, il faut corriger les nombres bruts d_0 et l_0 avant d'appliquer la formule (II 6).

Remarque. — Des observations analogues s'appliquent à plusieurs autres coefficients statistiques. Les uns peuvent présenter le caractère de probabilités mathématiques, si l'on élimine l'influence numérique des causes perturbatrices dont les principales sont les migrations, la natalité, la mortalité. D'autres coefficients ne présentent absolument pas le caractère de probabilité mathématique. Tel est par exemple le quotient $N : P$, où le dividende N représente le nombre des naissances dans un pays déterminé au cours d'une année X, et où le diviseur P est l'effectif de la population totale du dit pays au premier janvier de la même année ; manifestement, la condition formelle énoncée au début de ce numéro ne saurait être remplie dans ce cas. Ces quotients démographiques, prédilection de certains statisticiens, présentent de l'intérêt et peuvent être utiles ; mais il ne faudrait pas leur appliquer les théorèmes du calcul des probabilités.

30. — Il suit des considérations précédentes que la probabilité mathématique est toujours un nombre compris entre zéro et l'unité :

$$0 \leqslant p(E) = \frac{f}{s} \leqslant 1 \qquad \text{(II 7)}$$

Le nombre 1 représente la certitude, car l'événement E se réalisera sûrement si tous les cas lui sont favorables, et $f = s$ entraîne $p(E) = 1$ (v. n° 22, *d*). *Le nombre zéro représente l'impossibilité ;* en effet, la réalisation de l'événement E est impossible si aucun cas ne lui est favorable, et $f = 0$ entraîne $p(E) = 0$. En réalité, quand $p = 0$ ou $p = 1$, il y a certitude. Il ne s'agit donc plus à proprement parler de probabilité, puisque ni l'un ni l'autre de ces deux cas extrêmes n'implique le moindre doute. C'est pour éviter des exceptions de langage que les mathématiciens parlent d'une « probabilité nulle » ou de la « probabilité $p = 1$ » ; mais c'est une pure convention, car il y a une différence essentielle entre les deux états d'esprit (v. cependant n° 190, remarque 1).

On a essayé, en considérant des probabilités infiniment petites, c'est-à-dire variables et tendant vers zéro, de prouver qu'un événement physiquement impossible est celui dont la probabilité mathématique est infiniment petite. Il faut cependant établir une différence essentielle entre l'impossibilité et une probabilité qui tend vers zéro, de même entre la certitude et une probabilité qui tend vers l'unité. Aucun pont ne saurait être jeté sur cet abîme ni établir une transition de l'un à l'autre.

Quelque chose d'analogue se présente lorsqu'on envisage une suite convergente de nombres rationnels ; la limite vers laquelle ces nombres convergent peut fort bien être tout autre chose qu'un nombre rationnel. De même dans le domaine des fonctions : la limite vers laquelle tend une suite infinie convergente de fonctions toutes continues peut fort bien être une fonction discontinue ; etc.

Remarque. — Dans le langage courant, on dit souvent que « la probabilité d'un événement E est très grande », ce qui est une façon de parler incorrecte, puisque $p(E)$ ne dépasse jamais 1 ; on devrait donc dire que la probabilité est très voisine de l'unité. En réalité, c'est le rapport $p(E) : p(F)$ qui est très grand, donc une probabilité relative, savoir la probabilité de E comparée à celle de l'événement contraire F et mesurée en prenant cette dernière comme unité.

II. — *Les théorèmes fondamentaux.*

31. — Pour déterminer la probabilité mathématique d'un événement aléatoire E, la *méthode directe* consiste à établir des cas simples équipossibles, puis à déterminer les nombres s et f et à faire le rapport $f : s$ (v. n° 23). Dans le premier stade, cette méthode fut la seule employée pour déterminer $p(E)$. Peu à peu cependant, grâce à la tendance des mathématiques de ramener la solution de problèmes à celle d'autres moins compliqués, on parvint à établir des règles particulières dont la combinaison permettait de résoudre des problèmes où la méthode directe conduisait à des complications inextricables. Ces règles, d'abord assez nombreuses, se réduisirent finalement à un petit nombre de propositions fondamentales, grâce surtout à Laplace qui les condensa et le premier les énonça d'une façon précise dans son *Essai philosophique sur les probabilités* (v. n° 11). On leur donne parfois le nom de *méthodes indirectes* de détermination de la probabilité. Nous en démontrerons les deux plus importantes, après avoir rappelé quelques définitions.

Définition 5. — Un événement E est dit *composé* des deux événements E_1 et E_2, si E consiste dans l'arrivée simultanée ou successive de E_1 et de E_2. Dans ce cas, j'écrirai $E = E_1 E_2$ et désignerai la probabilité mathématique de cet événement composé par $p(E_1 E_2)$. De même, un événement composé de n autres sera désigné par $E_1 E_2 E_3 \ldots\ldots E_n$ et sa probabilité mathématique par $p(E) = p(E_1 E_2 E_3 \ldots\ldots E_n)$. L'événement composé E n'est donc censé réalisé que lorsque les n événements « partiels » E_λ se sont tous produits, soit successivement soit simultanément. Des événements peuvent être dépendants ou indépendants, compatibles ou exclusifs l'un de l'autre.

Définition 6. — Deux événements sont dits *indépendants*, si l'arrivée de l'un n'influence en rien la probabilité d'arrivée de l'autre. Si au contraire la réalisation du premier événement exerce quelque influence sur la probabilité mathématique du second, les événements sont dits *dépendants*. Cette définition s'étend aisément

au cas d'un nombre quelconque d'événements. Lorsqu'il y a dépendance entre deux événements A et B, au sens qu'on vient de définir, elle est en général unilatérale, A étant antérieur à B.

Définition 7. — Des événements sont dits *incompatibles* ou *exclusifs l'un de l'autre*, quand la réalisation de l'un quelconque d'entre eux exclut l'arrivée de tous les autres. Ils sont dits *compatibles* dans le cas contraire.

32. — *Le théorème d'addition du calcul des probabilités.* On l'appelle aussi « le théorème des probabilités totales », ou la probabilité du « soit-l'un soit-l'autre ». Il s'applique lorsqu'un événement E peut se réaliser de plusieurs manières différentes exclusives l'une de l'autre. α) *Quand un événement aléatoire E peut se réaliser de plusieurs manières s'excluant mutuellement, la probabilité mathématique de E est la somme des probabilités mathématiques qui correspondent aux divers modes de réalisation.*

Démonstration. — Supposons le phénomène ramené au schéma de l'urne. Envisageons une urne $U(a, b; s)$ contenant a boules blanches, b noires, au total $a + b = s$ boules. Supposons, en outre, chaque boule caractérisée par un numéro, soit 1 ou 2 ou 3 etc., ou n. Mettons qu'il y ait a_1 boules blanches portant le numéro 1 ; a_2 boules blanches portant le n° 2 ; etc., enfin a_n boules blanches portant le numéro n ; on a manifestement

$$a_1 + a_2 + \ldots\ldots + a_h + \ldots\ldots + a_n = a,$$

puisque le nombre total des boules blanches est a. Si l'on range en un même groupe les tirages qui amènent des boules portant le même numéro, la totalité des tirages possibles se trouvera décomposée en n groupes exclusifs les uns des autres. Divisant l'égalité précédente par s, on a

$$\frac{a_1}{s} + \frac{a_2}{s} + \ldots\ldots + \frac{a_h}{s} + \ldots\ldots + \frac{a_n}{s} = \frac{a}{s}.$$

Désignons par E la sortie d'une boule blanche sans égard au numéro qu'elle porte, par E_1 la sortie d'une boule blanche portant le numéro 1, par E_2 la sortie d'une boule blanche portant le

numéro 2, et ainsi de suite, enfin par E_n la sortie d'une boule blanche portant le numéro n. On a, en vertu des définitions posées (v. n° 23),

$$p(E) = \frac{a}{s} ; p(E_1) = \frac{a_1}{s} ; p(E_2) = \frac{a_2}{s} ; \ldots\ldots ; p(E_n) = \frac{a_n}{s} \qquad \text{(II 8)}$$

et la formule précédente devient :

$$p(E) = p(E_1) + p(E_2) + \ldots\ldots + p(E_n), \qquad \text{(II 9)}$$

formule qui exprime le théorème d'addition.

β) L'application de ce théorème présuppose les conditions suivantes :

1] L'événement étudié E doit pouvoir se réaliser de n manières différentes, savoir par l'arrivée de n événements dits partiels : E_1, $E_2, \ldots\ldots, E_n$, dont chacun ait une probabilité mathématique déterminée, $p(E_\lambda)$, dite *probabilité partielle* par opposition à $p(E)$ dite *probabilité totale*.

2] Ces n événements partiels E_λ doivent être exclusifs l'un de l'autre, au sens de la définition 7 ci-dessus.

3] L'arrivée de l'un quelconque des événements E_λ doit entraîner la réalisation de l'événement étudié E.

Des conditions analogues doivent se retrouver dans les cas favorables, savoir :

4] L'ensemble des cas favorables à E doit se départager en n ensembles partiels dont chacun se rapporte à l'un des événements partiels E_λ.

5] Ces n ensembles ne doivent avoir aucun élément commun.

Pour des exemples d'application du théorème des probabilités totales, voir les numéros 38, problème 14 ; n° 39 ; n° 61, remarque ; n° 62, remarque 2 ; n° 63 etc.

33. — *Le théorème de multiplication du calcul des probabilités.* On l'appelle aussi « le théorème des probabilités composées » ; ou la probabilité du « et-l'un-et-l'autre ». Il s'applique lorsque l'événement étudié consiste dans la production, simultanée ou successive, de plusieurs autres événements dits *simples* ou *compo-*

sants (v. n° 31, définition 5). Dans l'évaluation de la probabilité mathématique d'un événement composé $(E_1E_2E_3\ldots\ldots E_n)$, il y a lieu de distinguer si les événements composants E_λ sont, ou non, indépendants les uns des autres (v. n° 31, définition 6). Nous allons traiter séparément ces deux cas. Pour abréger le langage et l'écriture, convenons de ce qui suit.

Définition 8. — Le symbole $U(a, b\,;\, s)$ représente une urne U contenant a boules blanches, b noires, au total $s = a + b$ boules. De plus, à chaque tirage, on remet dans l'urne la boule tirée, après en avoir noté la couleur, et l'on mélange soigneusement à nouveau le contenu de l'urne. Enfin, pour réaliser au mieux les cas équipossibles, nous supposerons que les boules ne diffèrent les unes des autres que par la couleur et pour éliminer l'influence de celle-ci, nous imaginerons les tirages effectués par un aveugle.

Premier cas. — Les événements composants $E_1, E_2, \ldots\ldots, E_n$ sont indépendants au sens ci-dessus défini.

Commençons par le cas de 2 événements composants et envisageons deux urnes : $U_1(a_1, b_1\,;\, s_1)$ et $U_2(a_2, b_2\,;\, s_2)$.

Problème 7. — Un aveugle tire au hasard une boule de chacune de ces deux urnes. Quelle est la probabilité mathématique qu'il en sorte deux boules blanches? *Solution.* Que les tirages soient simultanés ou successifs, on voit que cela n'influe aucunement sur les probabilités mathématiques. Déterminons d'abord le nombre s de tous les cas possibles. A cet effet, imaginons que les boules de U_1 soient numérotées 1, 2, 3, ..., s_1 et les boules de U_2 numérotées 1, 2, 3, ..., s_2, afin de les distinguer entre elles. Si notre aveugle saisit d'une main la boule n° 1 de U_1 par exemple, il pourra sortir de l'autre urne, U_2, l'une quelconque des s_2 boules qui s'y trouvent. Cela fait s_2 cas équipossibles. Les mêmes s_2 possibilités se présentent si notre aveugle saisit dans U_1 la boule numéro 2 ou la boule numéro 3 etc. ou enfin la boule numéro s_1 ; à chacune des s_1 boules de la première urne correspondent ainsi s_2 cas équipossibles ; il y en a donc en tout $s_1.s_2$, de sorte que $s = s_1.s_2$; en ne considérant que les boules blanches, on démon-

trerait de même que le nombre des cas favorables est $f = a_1 . a_2$; la probabilité mathématique est donc

$$p(E) = \frac{f}{s} = \frac{a_1 a_2}{s_1 s_2} = \frac{a_1}{s_1} \cdot \frac{a_2}{s_2} .$$

Or, $a_1 : s_1$ est la probabilité mathématique de sortir boule blanche de U_1, donc $p(E_1)$; de même $a_2 : s_2 = p(E_2)$ = probabilité mathématique de tirer une boule blanche de U_2. Il vient donc

$$p(E) = p(E_1 E_2) = p(E_1) . p(E_2) \qquad \text{(II 10)}$$

Cette formule exprime le théorème de multiplication que l'on peut énoncer comme suit. *La probabilité mathématique d'un événement composé de deux autres est égale au produit des probabilités mathématiques des événements qui le composent.*

34. — Supposons trois événements composants et que

$$E = (E_1 E_2 E_3).$$

On envisagera l'arrivée des deux premiers comme un événement composé $A \equiv (E_1 E_2)$; alors $E = (A E_3)$ et l'on trouve, en appliquant deux fois le théorème précédent,

$$p(E) = p(A) . p(E_3) = p(E_1) . p(E_2) . p(E_3).$$

On procéderait de même, s'il y avait n événements composants. Par la conclusion de n à $n + 1$, on démontre facilement que le théorème reste vrai pour un nombre fini quelconque d'événements composants, pourvu qu'ils soient indépendants les uns des autres. Formule :

$$p(E_1 E_2 E_3 \ldots\ldots E_n) = p(E_1) . p(E_2) . p(E_3) \ldots\ldots p(E_n) \qquad \text{(II 11)}$$

Elle exprime le *théorème fondamental : La probabilité mathématique que plusieurs événements indépendants les uns des autres se réalisent tous, simultanément ou successivement, est égale au produit des probabilités mathématiques correspondantes.*

Corollaire. — Supposons en particulier

$$E_1 = E_2 = E_3 = \ldots\ldots = E_n = A.$$

L'événement étudié, E, consistera dans la réalisation n fois répétée de l'événement A. Je désignerai par analogie cet événement composé E par A^n ; la formule précédente devient

$$p(A^n) = p(A).p(A).p(A)..... \; p(A) = [p(A)]^n$$

ou plus simplement

$$p(A^n) = p^n(A), \qquad \text{(II 12)}$$

ce qui peut s'énoncer comme suit sous forme de *théorème. La probabilité mathématique qu'un événement* A *se répète n fois de suite dans les mêmes conditions est égale à la n*^{ième} *puissance de la probabilité mathématique de l'événement* A *lui-même.*

Remarque. - Si les événements composants, quoiqu'étant différents, ont la même probabilité mathématique, le théorème reste encore vrai. Exemple : La probabilité mathématique d'amener le six en lançant un dé à jouer parfait est $p = 1/6$ (v. n° 23, problème 1). La probabilité mathématique d'amener 6, 6, 6, 6, en quatre coups, est

$$\left(\frac{1}{6}\right)^4 = \frac{1}{1296} = 0{,}0007716.....$$

Et la probabilité mathématique d'amener en quatre coups 4, 3, 2, 1, dans cet ordre-là, est aussi 1 : 1296 ; celle d'amener 2, 5, 6, 3, dans cet ordre de succession, est encore 1 : 1296 et ainsi de suite. On croit communément que l'apparition de 6, 6, 6, 6 est moins probable que celle de 2, 5, 6, 3 par exemple ; c'est une erreur, provenant de ce que cette dernière combinaison, étant quelconque, ne frappe pas autant l'esprit. Ajoutons qu'il en serait autrement si l'on ne prescrivait pas l'ordre de succession des points 2, 5, 6, 3 ou qu'on lançât simultanément quatre dés (v. n° 38, problème 14).

35. — La probabilité mathématique étant toujours une fraction proprement dite, le théorème précédent prouve ceci : *La probabilité mathématique qu'un événement aléatoire E se répète* n *fois de suite diminue à mesure que* n *augmente ; elle tend vers zéro, quand* n *croît au delà de toute limite.* Dès lors, la probabilité

pour que l'événement contraire, F, finisse par se réaliser au moins une fois devient de plus en plus grande, au fur et à mesure qu'on augmente le nombre des épreuves. Appliquons cette remarque à un problème d'un grand intérêt historique, puisqu'il donna lieu à l'invention de la Théorie des chances.

Problème 8. — Quelle est la probabilité mathématique P pour qu'un événement E, de probabilité mathématique $= p$, se réalise *au moins une fois* dans une série de s épreuves ? *Solution.* Comme $p(E) = p$, la probabilité mathématique de l'événement contraire, F, est égale à $1 - p$; donc, la probabilité mathématique que F arrive n fois de suite est $(1 - p)^n$. Or, ceci est le contraire de l'événement étudié, car si F n'arrive pas n fois de suite dans une série de n épreuves, E se réalisera au moins une fois ; la probabilité cherchée P est donc

$$P = 1 - (1 - p)^n \qquad \text{(II 13)}$$

Il en résulte bien que $P \to 1$ quand $n \to \infty$, puisque $(1 - p)$ est compris entre zéro et 1 et que dès lors $(1 - p)^n \to 0$ quand $n \to \infty$. En d'autres termes, on a le

Corollaire : quelqu'invraisemblable que soit un événement E, quelque petite que soit sa probabilité mathématique p, *la réalisation de E devient presque certaine, si les épreuves deviennent suffisamment nombreuses.*

Application. — Si l'on se donne d'avance P, on peut tirer de (II 13) le nombre des épreuves nécessaires pour pouvoir, avec la probabilité mathématique P, s'attendre à ce que l'événement considéré, de probabilité $= p$, se réalise au moins une fois. Le problème implique $P > p$. La formule (II 13) donne en effet $(1 - p)^n = 1 - P$, d'où, en prenant les logarithmes,

$$n = \frac{\log (1 - P)}{\log (1 - p)} . \qquad \text{(II 14)}$$

Problème 9. — Combien de fois faut-il lancer un dé parfait pour pouvoir parier 1 contre 1 d'amener au moins une fois le six ? *Solution.* La probabilité mathématique d'amener le six en 1 coup

est $p = \frac{1}{6}$ (v. n° 23, problème 1), et « parier 1 contre 1 » équivaut à la probabilité $P = \frac{1}{2}$. La formule (II 14) donne donc

$$n = \log \frac{1}{2} : \log 5/6 = \frac{\log 2}{\log 6 - \log 5} = \frac{0{,}30103}{0{,}07918} = 3{,}801 \ldots$$

Comme n doit être un nombre entier, on prendra $n = 3$ ou $n = 4$; ce dernier chiffre étant supérieur au nombre trouvé 3,8....., on a avantage à parier 1 contre 1 d'amener le six au moins 1 fois en quatre coups, en lançant 1 dé. On le vérifie immédiatement, en résolvant le

Problème 10. — Quelle est la probabilité mathématique de ne jamais amener le point six, en lançant un dé à jouer parfait 4 fois de suite ? *Solution.* La probabilité mathématique de ne pas amener le six en 1 coup est $p = 5/6$; donc, celle de ne l'amener ni au premier ni au second ni au troisième ni au quatrième coup est $= (5/6)^4 = 625 : 1296 = 0{,}482\,253$; donc, la probabilité mathématique de l'événement contraire, qui consiste à amener le six au moins une fois en 4 coups, est $1 - 0{,}482\,253 = 0{,}517\,747$. Elle est effectivement supérieure à 1/2, et l'on peut parier 67 contre 62, ou 671 contre 625, d'amener le six au moins une fois en 4 coups.

36. — *Problème 11.* — Au jeu de trictrac, où l'on jette simultanément 2 dés, combien de coups faut-il jouer pour pouvoir parier 1 contre 1 d'amener au moins une fois le sonnez (v. n° 4, α)? *Solution.* La probabilité mathématique d'amener le sonnez en 1 coup est $p = 1/36$ (v. n° 23, problème 2) ; comme au problème précédent, nous avons ici $P = \frac{1}{2}$. La formule (II 14) devient dès lors :

$$n = \log \frac{1}{2} : \log 35/36 = \frac{\log 2}{\log 36 - \log 35} = \frac{0{,}30103}{0{,}01223} = 24{,}614 \ldots$$

Comme n doit être un nombre entier, on prendra $n = 24$ ou $n = 25$. Or, 24 étant inférieur au chiffre trouvé 24,614, ..., il y a désavantage à parier 1 contre 1 d'amener le sonnez au moins une fois en 24 coups.

Remarque historique. — Ce problème a un intérêt historique parce qu'il donne la réponse à l'une des questions posées à Pascal par le chevalier de Méré et se trouve ainsi lié à l'origine de la Théorie des chances (v. n° 4). Pascal écrivit à Fermat (lettre du 29 juillet 1654) : « Je n'ai pas le temps de vous envoyer la démonstration d'une difficulté qui étonnait fort M. de Méré, car il a un très bon esprit, mais il n'est pas géomètre; c'est, comme vous savez, un grand défaut. Il me disait donc qu'il avait trouvé fausseté dans les nombres pour cette raison : Si l'on entreprend de faire 6 avec un dé, il y a avantage de l'entreprendre par 4 coups ; si l'on entreprend de faire sonnez (double six) avec deux dés, il y a désavantage de l'entreprendre en 24 coups; et néanmoins 24 est à 36, qui est le nombre des faces de deux dés, comme 4 est à 6, qui est le nombre des faces d'un dé. Voilà quel était son grand scandale et qui lui faisait dire hautement que les propositions n'étaient pas constantes et que l'Arithmétique se démentait. Mais vous en verrez bien aisément la raison, par les principes où vous êtes ».

De Méré croyait donc que le nombre des coups nécessaires pour avoir le même avantage à parier 1 contre 1 doit être proportionnel au nombre des coups possibles, dans le cas particulier : égal aux deux tiers de ce nombre. Le calcul ci-dessus montre la fausseté de ce raisonnement. La différence devient encore plus sensible, si l'on prend un plus grand nombre de dés. Les problèmes suivants le prouvent.

Problème 12. — Combien de fois faut-il lancer simultanément 3 dés parfaits pour pouvoir parier 1 contre 1 d'amener au moins une fois « rafle de six », c'est-à-dire le triple six ? *Solution.* La probabilité mathématique d'amener le triple six au premier coup est $p = 1 : 216$, puisque 3 dés présentent $s = 6^3 = 216$ combinaisons possibles dont une seule réalise rafle de six (v. n° 23, problème 2). La formule (II 14) donne

$$n = \log \frac{1}{2} : \log \frac{215}{216} = \frac{\log 2}{\log 216 - \log 215} = \frac{0{,}30103}{0{,}00201} = 149{,}76\ldots$$

On prendra donc $n = 149$ ou $n = 150$. Or, les 2/3 des 216 coups possibles ne font que 144. Il y a avantage à parier 1 contre 1

d'amener, avec 3 dés, rafle de six en 150 coups, désavantage à le parier en 149 coups, désavantage encore plus prononcé à l'entreprendre en 144 coups.

Avec 4 dés lancés simultanément, il y a $s = 6^4 = 1296$ coups possibles dont les 2/3 font 864 ; mais la formule (II 14) donne dans ce cas $n = \frac{\log 2}{\log 1296 - \log 1295} = \frac{0{,}30103}{0{,}00034} = 885{,}382\ldots$

On aurait donc avantage à parier 1 contre 1 d'amener quadruple 6 au moins 1 fois en 886 coups, désavantage à l'entreprendre en 885 coups, désavantage à fortiori en 864 coups.

37. — *Deuxième cas.* — Les événements composants ne sont pas indépendants les uns des autres (v. n° 33). Pour adapter le schéma de l'urne à ce cas, reprenons la convention du n° 33, définition 8. Imaginons que le premier tirage se fasse de l'urne $U(a_1, b_1 ; s_1)$; que pour le deuxième tirage, on ait deux urnes à disposition : $U_e(a_1', b_1' ; s_1')$ et $U_f(a_2', b_2' ; s_2')$; que pour le troisième tirage, on ait 4 urnes à disposition, savoir 1) : $U_{ee}(a_1'', b_1'' ; s_1'')$, 2) : $U_{ef}(a_2'', b_2'' ; s_2'')$, 3) : $U_{fe}(a_3'', b_3'' ; s_3'')$ et 4) : $U_{ff}(a_4'', b_4'' ; s_4'')$; que pour le quatrième tirage, on ait 8 urnes à disposition : U_{eee}, U_{eef}, U_{efe}, ..., U_{fff} ; en général, supposons que pour le $n^{ième}$ tirage, on ait à disposition $2^n - 1$ urnes différentes, de compositions connues (v. n° 45).

Convenons enfin de la règle suivante : Le second tirage se fera de l'urne U_e ou de l'urne U_f, suivant que le premier aura fourni boule blanche ou boule noire. Le troisième tirage se fera de l'urne U_{ee} si les deux premiers auront amené boule blanche ; il se fera de l'urne U_{ef} si le premier tirage aura amené boule blanche et le second boule noire ; et ainsi de suite, les indices e ou f correspondant respectivement à l'apparition de boule blanche ou de boule noire dans les tirages précédents. Par exemple, le quatrième tirage s'opérera de l'urne U_{fee} si le premier avait amené boule noire et les deux autres boule blanche. A un tirage quelconque, la probabilité mathématique de la sortie de boule blanche dépendra donc des résultats de tous les événements composants qui l'auront précédé, et ceux-ci seront bien dépendants les uns des autres, au sens du n° 31.

Problème 13. — Quelle est la probabilité mathématique de tirer deux fois de suite boule blanche, dans les conditions spéciales susmentionnées? *Solution.* Le nombre s des cas possibles est $s_1 s_1'$ (v. n° 33, problème 7). Le nombre f des cas favorables à l'événement étudié est $a_1 a_1'$, puisque par convention, boule blanche étant sortie au premier coup, il faudra plonger la main dans $U_e(a_1', b_1'; s_1')$. On a donc

$$p(E) = \frac{a_1 a_1'}{s_1 s_1'} = \frac{a_1}{s_1} \cdot \frac{a_1'}{s_1'}.$$

Or, la probabilité mathématique de sortir boule blanche au premier tirage est $p(E_1) = a_1 : s_1$; la probabilité de sortir boule blanche au second tirage, *quand le premier a donné boule blanche*, est $p(E_2) = a_1' : s_1'$. On voit ainsi par la formule précédente que

$$p(E) = p(E_1) \cdot p(E_2), \qquad \text{(II 15)}$$

en paroles : *La probabilité mathématique d'un événement composé de deux autres dont le second dépend du premier et n'arrive que si le premier s'est déjà réalisé, est égale au produit des probabilités mathématiques des deux événements composants, en prenant comme probabilité mathématique du second, celle qu'il a* après *la réalisation du premier.*

On voit facilement comment généraliser le raisonnement au cas d'un nombre quelconque n d'événements composants, $E_1, E_2, \ldots\ldots, E_n$. On arrive à la formule

$$p(E_1 E_2 E_3 \ldots\ldots E_n) = p(E_1) \cdot p(E_2) \cdot p(E_3) \ldots\ldots p(E_n) \qquad \text{(II 16)}$$

qui est extérieurement identique avec la formule (II 11) du n° 34. Mais dans la formule (II 16), il faut entendre par $p(E_\lambda)$ la probabilité mathématique qu'a l'événement E_λ quand on sait que les événements $E_1, E_2, \ldots\ldots, E_{\lambda-1}$ qui le précèdent sont déjà tous arrivés et ont exercé leur influence sur $p(E_\lambda)$. En d'autres termes, dans la formule (II 16), $p(E_3)$ représente la probabilité mathématique qu'a E_3 *après* les arrivées de E_1 et de E_2; de même $p(E_4)$ est la probabilité mathématique qu'a E_4 *après* les arrivées de E_1, de E_2 et de E_3; et ainsi de suite jusqu'à E_n. On voit immédiatement que

cette restriction devient inutile dans le cas où les événements composants sont indépendants les uns des autres.

C'est presque toujours en combinant les théorèmes d'addition et de multiplication des probabilités qu'on arrive le plus rapidement à la solution des problèmes posés. Nous allons le montrer sur quelques exemples utiles pour la suite.

38. — *Le jeu de pile ou face* est le plus simple possible des jeux envisagés par la Théorie des chances. On jette en l'air une pièce de monnaie ; l'un des côtés s'appelle pile, l'autre face. On engage des paris sur le côté qui sera apparent après la chute. Pour pouvoir appliquer la définition classique (v. les nos 19, 23 et 27), il faut faire l'hypothèse que la pièce *est bonne pour ce jeu*, c'est-à-dire que les chances d'apparition sont égales pour le côté pile et le côté face. Cela implique que la pièce soit physiquement homogène et d'une symétrie parfaite au point de vue géométrique, de sorte que son centre de gravité coïncide exactement avec son centre de figure. Si l'on pense que le relief irrégulier de la pièce de monnaie ne permettrait jamais qu'elle soit rigoureusement « bonne », il faut supposer que l'on prenne un jeton fabriqué avec le plus grand soin et dont les deux côtés ne diffèrent par exemple que par la couleur d'une petite croix peinte en noir sur le côté appelé pile, et en rouge sur le côté appelé face (v. n° 23, remarque).

Problème 14. — Dans une série de 4 coups à pile ou face, quelle est la probabilité mathématique d'amener au moins deux fois face ? *Solution.* En vertu de l'hypothèse de l'équipossibilité des cas, la probabilité mathématique d'amener pile est $p = \frac{1}{2}$, celle d'amener face est $q = \frac{1}{2}$. L'événement $E \equiv$ apparition de face au moins deux fois peut, ici, se réaliser de trois façons ; E se produit en effet 1) quand face arrive 4 fois de suite, 2) quand face se présente 3 fois et pile 1 fois, 3) quand face se montre 2 fois et pile 2 fois. Comme ces trois modalités sont exclusives l'une de l'autre, on peut appliquer le théorème d'addition (v. n° 32, β). Commençons dès lors par déterminer la probabilité mathématique de chacune de ces modalités.

1) La probabilité mathématique d'amener face 4 fois de suite est $\left(\frac{1}{2}\right)^4 = \frac{1}{16}$, en vertu du théorème des probabilités composées (v. n° 34, corollaire).

2) La probabilité mathématique d'amener face 3 fois et pile 1 fois, *dans un ordre de succession déterminé*, par exemple *EEEF*, est $\left(\frac{1}{2}\right)^3 \cdot \left(\frac{1}{2}\right) = \frac{1}{16}$. Or, il y a 4 ordres de succession possibles, savoir

EEEF, *EEFE*, *EFEE*, *FEEE*.

Chacun d'eux a la même probabilité mathématique de se réaliser ; et comme ils sont exclusifs l'un de l'autre, on peut appliquer le théorème d'addition. Donc, la probabilité mathématique d'amener face 3 fois et pile 1 fois, *dans un ordre de succession quelconque*, est

$$4 \cdot \left(\frac{1}{2}\right)^3 \cdot \frac{1}{2} = \frac{4}{16} = \frac{1}{4}.$$

3) La probabilité mathématique d'amener face 2 fois et pile 2 fois, *dans un ordre déterminé*, par exemple *EEFF*, est $\left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^2$, en vertu du théorème de multiplication (v. n° 34). Or, il y a 6 ordres de succession possibles, savoir

E^2F^2, *EFEF*, EF^2E, FE^2F, *FEFE*, F^2E^2.

Chacun de ces 6 arrangements est également enclin à se réaliser, vu l'hypothèse de la pièce « bonne » ; et comme ils sont exclusifs l'un de l'autre, on peut appliquer de nouveau le théorème d'addition. La probabilité mathématique d'amener face 2 fois et pile 2 fois, *dans un ordre de succession quelconque*, est donc

$$6 \cdot \left(\frac{1}{2}\right)^2 \cdot \left(\frac{1}{2}\right)^2 = \frac{6}{16} = \frac{3}{8}.$$

Il ne reste plus qu'à appliquer le théorème d'addition. On trouve :

$$\frac{1}{16} + \frac{4}{16} + \frac{6}{16} = \frac{11}{16} = 0{,}6875.$$

Réponse. — La probabilité mathématique d'amener face au

moins deux fois en quatre coups est 11 : 16, et l'on peut parier 11 contre 5 de le faire dans un ordre de succession quelconque non prescrit d'avance.

39. — La *loterie génoise* ou le *lotto des nombres. Problème 15.* — Quelle est la probabilité mathématique de gagner à « l'extrait simple » dans la loterie génoise? Voici les conditions de ce problème. Une urne contient 90 billets numérotés de 1 à 90. On y fait une série de $r = 5$ tirages consécutifs, *sans* remettre le billet sorti dans l'urne. On demande la probabilité mathématique qu'un numéro choisi d'avance entre 1 et 90, au hasard, mettons le numéro 19, se trouvera parmi les cinq numéros tirés. *Solution.* L'événement E, sortie du numéro 19, peut ici se réaliser de cinq manières différentes : le dit numéro peut être tiré au premier coup ou au second coup ou au troisième ou au quatrième ou enfin au cinquième coup. Soient p_1, p_2, p_3, p_4, p_5, les probabilités mathématiques correspondantes. Ces cinq modalités étant exclusives l'une de l'autre, on peut appliquer le théorème d'addition (v. n° 32) ; la probabilité demandée est donc

$$p(E) = p_1 + p_2 + p_3 + p_4 + p_5.$$

Il reste à déterminer ces cinq probabilités partielles.

a) La probabilité mathématique de tirer le numéro 19 au premier coup est $p_1 = 1 : 90$ (v. n° 23).

b) La probabilité p_2 est celle d'un événement composé de deux autres ; il faut en effet : 1° que le numéro 19 ne sorte pas au premier coup, ce dont la probabilité mathématique est 89 : 90 ; et 2° que le dit numéro sorte au deuxième coup, ce dont la probabilité mathématique est 1 : 89, car il n'y a plus que 89 billets dans l'urne, puisque les billets tirés n'y sont pas remis. On peut donc appliquer le théorème des probabilités composées (v. n° 33), de sorte que

$$p_2 = \frac{89}{90} \cdot \frac{1}{89} = \frac{1}{90}.$$

c) La probabilité p_3 est celle d'un événement composé de trois autres ; il faut en effet : 1° que le numéro en question ne sorte pas

au premier coup, ce dont la probabilité mathématique est 89 : 90 ; 2° qu'il ne sorte pas non plus au second coup, ce dont la probabilité mathématique est = 88 : 89 ; et 3° que le dit numéro sorte au troisième coup, ce dont la probabilité mathématique = 1 : 88, puisque maintenant l'urne ne contient plus que 88 billets. Le théorème des probabilités composées donne dès lors

$$p_3 = \frac{89}{90} \cdot \frac{88}{89} \cdot \frac{1}{88} = \frac{1}{90} .$$

d) Par un raisonnement tout analogue, on démontrerait que

$$p_4 = \frac{89}{90} \cdot \frac{88}{89} \cdot \frac{87}{88} \cdot \frac{1}{87} = \frac{1}{90} , \qquad \text{et enfin}$$

e) que $$p_5 = \frac{89}{90} \cdot \frac{88}{89} \cdot \frac{87}{88} \cdot \frac{86}{87} \cdot \frac{1}{86} = \frac{1}{90} .$$

La probabilité mathématique cherchée, somme de ces cinq probabilités partielles, est donc

$$p(E) = \frac{5}{90} = \frac{1}{18} = 0{,}05555\ldots .$$

III. — *L'analyse combinatoire résumée.*

40. — Comme nous l'avons déjà fait remarquer, il y a connexion étroite entre le calcul des probabilités et l'analyse combinatoire (voir la fin du n° 5). Il est dès lors utile de résumer les formules fondamentales de cette analyse. Soient n objets ou choses quelconques, différents les uns des autres, concrets ou abstraits, dits *éléments*, par exemple des cartes, des chiffres, des lettres, des nombres, etc.

LES PERMUTATIONS

Définition 9. — Former les *permutations simples* ou « permutations sans répétition » de n éléments, c'est former, de toutes les manières possibles, des ensembles composés de ces mêmes éléments, mais pris dans différents ordres de succession. *Notation.* Désignons par *Perm* (n) le nombre total de ces permutations.

Avec 2 éléments, a et b, on ne peut former que 2 permutations, savoir ab et ba. Donc, $Perm\ (2) = 2$. Un troisième élément, c, peut se mettre à 3 places différentes dans chacune des permutations précédentes. On déduit en effet

de ab les groupes cab, acb, abc ;

de ba les groupes cba, bca, bac. Donc,

$$Perm\ (3) = 3.\ Perm\ (2) = 6.$$

En généralisant ce raisonnement, on voit sans peine que

$$Perm\ (4) = 4.\ Perm\ (3),$$
$$Perm\ (5) = 5.\ Perm\ (4),$$
$$\cdots\cdots\cdots$$
$$Perm\ (n) = n.\ Perm\ (n-1).$$

Multipliant ces égalités membre à membre, puis simplifiant, on trouve

$$Perm\ (n) = 1.2.3.4\ldots\ldots(n-1).n \equiv n! \qquad (II\ 17)$$

prononcé « n factorielle » ; les auteurs anglais désignent souvent ce produit des n premiers nombres naturels par $\lfloor\underline{n}$.

Exemple. — Combien de nombres à 9 chiffres peut-on former avec les 9 chiffres significatifs, chaque chiffre étant pris une seule fois ? Il s'agit manifestement des permutations de 9 éléments. Leur nombre est $9! = 1.2.3.4.5.6.7.8.9 = 362\ 880$.

Remarque. — On convient de poser $0! = 1$. Des considérations d'analyse supérieure justifient cette convention, et même la rendent nécessaire pour éviter des exceptions.

Les arrangements.

41. — *Définition 10.* — Les *arrangements simples* ou « arrangements sans répétition » de n éléments distincts pris p à p sont les groupes que l'on peut former en prenant, de toutes les manières possibles, p de ces éléments pour les ranger dans un ordre déterminé, un élément ne pouvant pas figurer plus d'une fois dans un même groupe. Des arrangements constitués des mêmes éléments, mais dans un ordre de succession différent, sont considérés comme

distincts. Deux arrangements peuvent donc différer l'un de l'autre soit par les éléments qui les composent, soit seulement par l'ordre de succession de ces éléments. *Exemples* d'arrangements : les mots que l'on peut former avec des lettres données ou les nombres que l'on peut écrire avec des chiffres donnés, car l'ordre de succession des lettres ou des chiffres n'est pas indifférent. *Notation.* Désignons par $Ar_p(n)$ le nombre de tous les arrangements simples possibles p à p de n éléments distincts.

Soient n lettres distinctes : $a, b, c, \ldots\ldots, l$. En les prenant une à une, on peut évidemment former n groupes distincts, de sorte que

$$Ar_1(n) = n.$$

Pour en déduire les arrangements 2 à 2, il faut ajouter à chacun de ces groupes d'une lettre les $(n - 1)$ autres. On déduit ainsi
de a les groupes $ab, ac, ad, \ldots\ldots, al$;
de b » $ba, bc, bd, \ldots\ldots, bl$;
de c » $ca, cb, cd, \ldots\ldots, cl$, et ainsi de suite.

De chacun des groupes primordiaux, on déduit $(n - 1)$ arrangements 2 à 2, de sorte que

$$Ar_2(n) = (n - 1).Ar_1(n).$$

On voit de même, sans difficulté, que

$$Ar_3(n) = (n - 2).Ar_2(n);$$
$$Ar_4(n) = (n - 3).Ar_3(n);$$
$$\ldots\ldots\ldots\ldots\ldots$$
$$Ar_p(n) = (n - p + 1).Ar_{p-1}(n).$$

Multipliant ces égalités membre à membre, puis simplifiant, on trouve

$$Ar_p(n) = n(n - 1)(n - 2)\ldots\ldots(n - p + 1) \qquad \text{(II 18)}$$

Exemple. — Combien de nombres à 5 chiffres peut-on former avec les 9 chiffres significatifs, chacun étant pris une seule fois ? Il s'agit évidemment de groupes ordonnés, puisque l'ordre de succession des chiffres n'est pas indifférent. Il faut donc trouver le nombre des arrangements simples de 9 éléments pris 5 à 5 ; c'est

$$Ar_5(9) = 9.8.7.6.5 = 36.42.10 = 15\,120.$$

Remarque 1. — Dans les arrangements simples, le nombre p des éléments qui entrent dans chaque groupe ne peut dépasser le nombre total des éléments disponibles, $p \leqslant n$. Dans le cas extrême où $p = n$, c'est-à-dire si l'on prend les n éléments n à n, on obtient des permutations. La formule (II 18) devient effectivement

$$Ar_n(n) = n(n-1)(n-2) \ldots\ldots 3.2.1 = n! \quad \text{(II 18}a\text{)}$$

Les permutations sont donc des cas particuliers d'arrangements. L'ordre de succession n'est pas indifférent.

Remarque 2. — La formule des arrangements simples peut s'écrire à l'aide de factorielles. On trouve en effet

$$Ar_p(n) = n! : (n-p)! \qquad \text{(II 19)}$$

Les combinaisons.

42. — *Définition 11.* — Les *combinaisons simples* ou « combinaisons sans répétition » de n éléments distincts pris p à p sont les groupes que l'on peut former en réunissant, de toutes les façons possibles, p de ces n éléments de manière qu'un élément ne figure pas plus d'une fois dans un même groupe. Des combinaisons constituées des mêmes éléments, mais dans un ordre de succession différent, ne sont pas considérées comme distinctes, contrairement à ce qui a lieu pour les arrangements. Deux combinaisons ne peuvent dès lors différer l'une de l'autre que par la nature des éléments qu'elles contiennent. *Exemples* de combinaisons : les produits que l'on peut former avec des facteurs donnés, car la valeur d'un produit est indépendante de l'ordre de succession de ses facteurs. *Notation.* Désignons par $C_p(n)$ le nombre des combinaisons simples possibles de n éléments distincts pris p à p.

Avec 4 nombres distincts, par exemple 2, 3, 5, 11, on peut former 4 produits différents de 3 facteurs, savoir 2.3.5, 2.3.11, 2.5.11, 3.5.11, de sorte que $C_3(4) = 4$. Chacun de ces 4 produits donne lieu à 6 arrangements distincts 3 à 3 ; par exemple, du premier produit dérivent

2.3.5, 2.5.3, 3.2.5, 3.5.2, 5.2.3, 5.3.2,

qui sont considérés, par définition, comme des arrangements distincts, mais non comme des combinaisons distinctes. Il y a donc 6 fois plus d'arrangements 3 à 3 que de combinaisons 3 à 3 ; en formule :

$$Ar_3(n) = 6 . C_3(n) = 3 ! \; C_3(n)$$

Pour la même raison, il y a 24 fois plus d'arrangements 4 à 4 que de combinaisons 4 à 4, ce que traduit la formule

$$Ar_4(n) = 24 . C_4(n) = 4 ! \; C_4(n).$$

D'une manière générale, les définitions 9 et 11 entraînent

$$Ar_p(n) = p ! \; C_p(n).$$

On en tire par division, en tenant compte de (II 17) et (II 18),

$$C_p(n) = \frac{n(n-1)(n-2) \ldots\ldots (n-p+1)}{1 . 2 . 3 \ldots\ldots p} \qquad \text{(II 20)}$$

Remarque. — On peut écrire (II 20) à l'aide de factorielles, en amplifiant la fraction par $(n - p) !$ Il vient

$$C_p(n) = \frac{n !}{p ! \, (n-p) !} \qquad \text{(II 21)}$$

43. — *Définition 12.* — On désigne généralement, depuis Léonard Euler (1707-1783), les expressions (II 20) et (II 21) par le symbole $\binom{n}{p}$, prononcé « n binomial p ». On appelle ce symbole « le coefficient binomial *de base* n et *d'indice* p ». On a donc, par cette définition,

$$\binom{n}{p} \equiv \frac{(n)(n-1)(n-2) \ldots\ldots (n-p+1)}{1 . 2 . 3 \ldots\ldots p} = \frac{n !}{p ! (n-p) !} \qquad \text{(II 22)}$$

Voici les propriétés les plus importantes des coefficients binomiaux.

a] Etant donnés n objets, quand on en prend 3 par exemple, on laisse un groupe de $(n - 3)$; à chaque combinaison 3 à 3 corres-

pond donc une combinaison $(n-3)$ à $(n-3)$; il y en a dès lors autant des unes que des autres, de sorte que

$$\binom{n}{3} = \binom{n}{n-3}, \text{ plus généralement } \binom{n}{p} = \binom{n}{n-p} \quad \text{(II 23)}$$

Donc : *deux coefficients binomiaux de même base ont la même valeur, si la somme de leurs indices est égale à la base commune.*

Par exemple $\binom{10}{7} = \binom{10}{3} \equiv \frac{10 \cdot 9 \cdot 8}{1 \cdot 2 \cdot 3} = 10 \cdot 3 \cdot 4 = 120.$

b] Il résulte directement de la définition 12 que l'on a toujours $\binom{n}{n} = 1$, donc aussi, en vertu de (II 23), que $\binom{n}{0} = 1$ (II 24)

c] Quand la base, n, est un nombre naturel, (II 22) et (II 23) entraînent

$$\binom{n}{-a} = \binom{n}{n+a} = 0 \quad \text{(II 25)}$$

d] Le symbole eulérien $\binom{x}{p}$ conserve sa signification, si la base x est quelconque. En particulier,

$$\binom{-1}{3} = \frac{(-1)(-2)(-3)}{1 \cdot 2 \cdot 3} = -1, \text{ et en général } \binom{-1}{n} = (-1)^n \quad \text{(II 26)}$$

$$\binom{3/2}{4} = \frac{3/2 \cdot 1/2 \cdot -1/2 \cdot -3/2}{1 \cdot 2 \cdot 3 \cdot 4} = \frac{3}{2.4.2^4} = \frac{3}{128}.$$

Remarque. — Le nombre des combinaisons simples de n éléments distincts pris p à p est ainsi $C_p(n) = \binom{n}{p}$. Dans le symbole eulérien, devenu international, c'est le chiffre supérieur, n, qui représente le nombre total des éléments. C'est pour cette raison que j'écris $C_p(n)$ et $Ar_p(n)$, bien que la plupart des auteurs écrivent C_n^p, A_n^p, quand le nombre des éléments est n.

e] Les coefficients binomiaux ont fait l'objet de nombreuses investigations, depuis que Pascal, dans son célèbre *Traité du triangle arithmétique* (11), les a étudiés (v. n° 10). Ils interviennent notamment dans la formule du binome de Newton, que nous

rappelons ici en vue des applications ultérieures et en appliquant les notations modernes.

$$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3 \text{ peut s'écrire}$$

$$= \binom{3}{0} a^3 + \binom{3}{1} a^2b + \binom{3}{2} ab^2 + \binom{3}{3} b^3 = \sum_{n=0}^{n=3} \binom{3}{n} . a^{3-n} . b^n .$$

De même,

$$(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4 \text{ peut s'écrire}$$

$$= \binom{4}{0} a^4 + \binom{4}{1} a^3b + \binom{4}{2} a^2b^2 + \binom{4}{3} ab^3 + \binom{4}{4} a^0b^4$$

$$= \sum_{n=0}^{n=4} \binom{4}{n} . a^{4-n} . b^n .$$

En général,

$$(a+b)^r = a^r + \frac{r}{1} a^{r-1} b + \frac{r(r-1)}{1.2} a^{r-2} b^2 + \ldots \ldots + b^r$$

$$= \binom{r}{0} a^r + \binom{r}{1} a^{r-1} b + \binom{r}{2} a^{r-2} b^2 + \ldots\ldots + \binom{r}{r} a^0 b^r$$

$$= \sum_{\lambda=0}^{r} \binom{r}{\lambda} a^{r-\lambda} . b^{\lambda} \qquad \text{(II 27)}$$

44. — *Permutations avec répétitions.* — Soient n éléments non tous distincts ; supposons que a d'entre eux soient identiques et égaux à l'objet A, que b autres soient identiques entre eux et égaux à l'objet B différent de A, et ainsi de suite jusqu'aux t derniers éléments supposés identiques entre eux, mais différents de tous les précédents. L'ensemble des n éléments donnés se décompose en groupes partiels tels que

$$n = a + b + c + \ldots\ldots + t.$$

Les permutations différentes que l'on peut former à l'aide de ces éléments sont dites « permutations avec répétition ». Leur nom-

bre, que je désigne par ${}^{r}Perm\ (n)$, est donné par la formule

$$ {}^{r}Perm\ (n) = \frac{n!}{a!\,b!\,c!\,\ldots\ldots\,t!} \qquad \text{(II 28)} $$

Démonstration. — Soient par exemple les 5 éléments a_1, a_2, a_3, b, c. Ils donnent lieu à 5 ! permutations. Celles-ci formées, qu'on efface les indices 1, 2, 3 ; on aura les éléments a, a, a, b, c ; et le nombre des groupes distincts sera manifestement devenu 3 ! fois plus petit, savoir 5 ! : 3 ! En généralisant ce raisonnement, on démontre la formule (II 28).

45. — *Arrangements avec répétitions. Définition 13.* — Les « arrangements avec répétitions de n éléments distincts pris p à p » sont les groupes que l'on peut former en prenant, de toutes les manières possibles, p de ces éléments, différents ou non, pour les disposer dans un ordre déterminé. Tout élément peut figurer jusqu'à p fois dans un groupe. Des arrangements constitués par les mêmes éléments mis dans des ordres de succession différents sont considérés comme distincts. Deux arrangements avec répétitions peuvent donc différer l'un de l'autre soit par les éléments qu'ils contiennent, soit seulement par l'ordre de succession de ces éléments ; v. n° 41, définition 10. En désignant le nombre des arrangements en question par ${}^{r}Ar_p(n)$, on a la formule

$$ {}^{r}Ar_p(n) = n^p. \qquad \text{(II 29)} $$

On voit en effet qu'en mettant 1 seul élément par groupe, on a

$$ {}^{r}Ar_1(n) = n. $$

Chacun donne lieu à n arrangements binaires :
de a on déduit $aa, ab, ac, \ldots, al$;
de b on déduit $ba, bb, bc, \ldots\ldots bl$; etc., de sorte que

$$ {}^{r}Ar_2(n) = n.{}^{r}Ar_1(n) ; $$

on verrait de même que

$$ {}^{r}Ar_3(n) = n.{}^{r}Ar_2(n) ; $$

$$ \ldots\ldots\ldots\ldots $$

$$ {}^{r}Ar_p(n) = n.{}^{r}Ar_{p-1}(n). $$

En multipliant ces égalités membre à membre, on trouve (II 29).

Exemple. — L'alphabet Morse ou alphabet des télégraphistes. Il ne comprend que deux éléments, le point . et le trait —; pris 1 à 1, ces éléments donnent 2 lettres, . pour e; — pour t; pris 2 à 2, ils donnent 4 arrangements, savoir .. pour i; .— pour a; — . pour n; — — pour m; pris 3 à 3, ils donnent lieu à $2^3 = 8$ arrangements, savoir

. . . ≡ s;	. . — ≡ u;	. — . ≡ r;	— . . ≡ d;
. — — ≡ w;	— . — ≡ k;	— — . ≡ g;	— — — ≡ o;

pris 4 à 4, ils donnent lieu à $2^4 = 16$ arrangements, savoir

. . . . ≡ h	— . . . ≡ b	— . . — ≡ x	— . — — ≡ y
. . . — ≡ v	. . — — ≡ $\ddot{u}$	— . — . ≡ c	— — . — ≡ q
. . — . ≡ f	. — . — ≡ $\ddot{a}$	— — . . ≡ z	— — — . ≡ $\ddot{o}$
. — . . ≡ l	. — — . ≡ p	. — — — ≡ j	— — — — ≡ ch

On y représente les chiffres par

. — — — — ≡ 1	— . . . . ≡ 6
. . — — — ≡ 2	— — . . . ≡ 7
. . . — — ≡ 3	— — — . . ≡ 8
. . . . — ≡ 4	— — — — . ≡ 9
. . . . . ≡ 5	— — — — — ≡ 0

46. — *Combinaisons avec répétitions. Définition 14.* — Les « combinaisons avec répétitions de n éléments distincts pris p à p » sont les groupes que l'on peut former en réunissant, de toutes les façons possibles, p des éléments donnés, différents ou non ; tout élément peut figurer jusqu'à p fois dans le même groupe. Des combinaisons qui ne diffèrent les unes des autres que par l'ordre de succession de leurs éléments ne sont pas considérées comme distinctes, contrairement à ce qui a lieu pour les arrangements ; v. n° 42, définition 11. Désignons par ${}^{r}C_{p}(n)$ le nombre des combinaisons en question.

Exemple. — Avec les 3 lettres a, b, c, on peut former : 6 combinaisons 2 à 2 avec répétitions, savoir aa, ab, ac, bb, bc, cc ; et

10 combinaisons 3 à 3 avec répétitions, savoir *aaa*, *aab*, *aac*, *abb*, *abc*, *acc*, *bbb*, *bbc*, *bcc*, *ccc*. Donc,

$${}^{r}C_2(3)=6\,;\quad {}^{r}C_3(3)=10\,;\quad {}^{r}C_4(3)=15\,;\quad {}^{r}C_5(3)=21\,;\ \text{etc.}$$

On remarque que

$$10=\frac{3+2}{3}\,.6\,;\ \text{que}\ 15=\frac{4+2}{4}\,.10\,;\ \text{que}\ 21=\frac{5+2}{5}\,.15\,;\ \text{etc.}$$

On établit la formule de récurrence générale suivante :

$${}^{r}C_s(n)=\frac{s+n-1}{s}\,.\,{}^{r}C_{s-1}(n)\,;$$

En y faisant successivement $s=1, 2, 3, 4, \ldots\ldots, p$, puis multipliant membre à membre les égalités ainsi obtenues, on en déduit :

$${}^{r}C_p(n)=\frac{n(n+1)(n+2)\ldots\ldots(n+p-1)}{1\,.\,2\,.\,3\ \ldots\ldots\ p}\qquad \text{(II 30)}$$

Remarque. — On peut aussi exprimer ce nombre à l'aide de factorielles, en amplifiant par $(n-1)!$; il vient :

$${}^{r}C_p(n)=\frac{(p+n-1)!}{p!\,(n-1)!}\,.\qquad \text{(II 31)}$$

IV. — *La formule de Stirling.*

47. — Les éléments de l'Analyse combinatoire sont résumés dans les sept formules (II 17), (II 19), (II 21), (II 22), (II 28), (II 29) et (II 31). Dans six d'entre elles figurent des factorielles. Aussi a-t-on reconnu de bonne heure l'utilité de posséder des tables de factorielles (44). Or, les factorielles croissent avec une rapidité extraordinaire. Déjà $20!=2\cdots 432\ 902\cdots 008\ 176\cdot 640\ 000$

(44) C. F. Degen a publié la première table de factorielles ; elle donne, pour $n=1, 2, 3, 4, \ldots\ldots$ jusqu'à $n=1200$, les logarithmes à 12 décimales de $n!$ *Tabularum ad faciliorem probabilitatis computationem utilem Enneas*, Copenhague, 1824.

s'écrit avec 19 et 30 ! avec 33 chiffres ; 400 ! s'écrit avec 869 chiffres, 800 ! avec 1977 chiffres ; 1200 ! a 3176 chiffres etc. Même écrits en caractères très petits, 2 millimètres par chiffre, 400 ! couvrirait une longueur d'environ 1,7 m. ; 800 ! une longueur de presque 4 mètres et 1200 ! une longueur de presque 6 $^1/_2$ mètres. Quand on a affaire à de grands nombres n, ce qui est le cas général dans les applications du calcul des probabilités, où n dépasse fréquemment 1000 et souvent 10 000, on est bien obligé de recourir à des formules plus expéditives, quand bien même elles ne donnent que des résultats approximatifs.

Observons qu'il n'est pas non plus nécessaire, dans les applications, d'avoir un résultat rigoureusement exact, le but étant de prévoir « à peu près » la fréquence relative des événements. Pour abréger les calculs, on se sert d'une formule découverte par Abraham de Moivre (v. n° 10, note 14), perfectionnée par Stirling ([45]). Chose curieuse, les nombres e, base des logarithmes népériens, et π, rapport de la circonférence au diamètre, entrent dans cette formule avec les valeurs bien connues $e = 2{,}7\ 1828\ 1828\ 459\ 045\ \ldots\ldots$, $\pi = 3{,}141\ 592\ 653\ 589\ 793\ 238\ 462\ 643\ 383\ 279\ \ldots\ldots$

Voici la formule en question, telle que la donna le grand Euler dans un mémoire de 1749, publié en 1753 ([45 a])

$$n! = \frac{n^n}{e^n}\sqrt{2\pi n}\cdot\left(1 + \frac{1}{12n} + \frac{1}{288n^2} - \frac{139}{51\,840n^3} + \ldots\ldots\right) \quad \text{(II 32)}$$

Dans la plupart des cas, on obtient une approximation déjà suffisante en conservant seulement le premier terme. La formule se réduit alors à

$$n! = \frac{n^n}{e^n}\cdot\sqrt{2\pi n}. \quad \text{(II 33)}$$

En prenant les logarithmes vulgaires des deux membres et introduisant le module $M \equiv \log e = 0{,}434\ 29\ \ldots\ldots$, on trouve

$$\log(n!) = \left(n + \frac{1}{2}\right)\cdot\log n - n\cdot 0{,}434\ 294\ 481\ 90\ 3\ 251\ 827\ 651\ldots\ldots + 0{,}399\ 089\ 9367\ \ldots\ldots \quad \text{(II 34)}$$

([45]) J. Stirling, *Methodus differentialis, sive Tractatus de summatione et interpolatione serierum infinitarum*, Londres, 1730.

([45 a]) Voir *Leonhardi Euleri Opera Omnia*, publiés par la Soc. Helvétique des Sciences Naturelles, série I, tome 14, page 515.

48. — Dans l'équation (II 33), le signe $=$ est employé très improprement, car les deux membres ne sont pas égaux entre eux, et leur différence ne tend pas non plus vers zéro, quand $n \to \infty$. Au contraire ! cette différence, qui représente *l'erreur absolue* commise en substituant $n^n \cdot \sqrt{2\pi n} : e^n$ à la place de $n!$, augmente avec n au delà de toute limite. C'est le rapport des deux membres qui tend vers 1, quand $n \to \infty$; pour être rigoureux, il faudrait donc écrire

$$n! = n^n \cdot \sqrt{2\pi n} \cdot e^{-n} \cdot (1 + \varepsilon_n) \qquad \text{(II 35)}$$

où ε_n représente un nombre positif qui $\to 0$ quand $n \to \infty$. On démontre que ε_n est à peu près $= 1 : 12n$, en tout cas inférieur à $1 : 11n$. Comme *l'erreur relative* provenant de l'emploi de (II 33) est égale à $\varepsilon_n : (1 + \varepsilon_n) = \varepsilon_n - \varepsilon_n^2 + \varepsilon_n^3 - \varepsilon_n^4 + \ldots$, on peut négliger les puissances supérieures, vu la petitesse de ε_n, et dire que l'erreur relative est $= \varepsilon_n$. On constate que

$\varepsilon_n < 0{,}0084$	pour	$n \geqslant 10$	$\varepsilon_5 = 0{,}016$
$\varepsilon_n < 0{,}0042$	«	$n \geqslant 20$	$\varepsilon_{10} = 0{,}0083$
$\varepsilon_n < 0{,}0028$	«	$n \geqslant 30$	$\varepsilon_{20} = 0{,}0041$
$\varepsilon_n < 0{,}00084$	«	$n \geqslant 100$	

Exemples. — 1) Pour $n = 20$, la formule (II 34) donne :

$$\log(20!) = 20{,}5 \cdot \log 20 - 20 \cdot 0{,}4342944819 + 0{,}3990899367$$
$$= 18{,}384315\,3 = \log 2\ 422\ 786\ 385\ 510\ 400\ 000$$

La valeur exacte est $20! = 2 \cdot\cdot\cdot 432\ 902 \cdot\cdot 008\ 176 \cdot 640\ 000$

Différence = erreur absolue = $10\ 115 \cdot\cdot 622\ 666 \cdot 240\ 000$

Cette erreur, très grande en valeur absolue, n'est cependant que la 243ième partie, soit environ $4{,}1\ {}^0/_{00}$, du nombre à évaluer.

2) Prenons $n = 30$; la valeur exacte est

$$30! = 265^{\text{V}}\ 252\ 859^{\text{IV}}\ 810\ 729^{\text{III}}\ 458\ 636^{\text{II}}\ 308\ 480^{\text{I}}\ 000\ 000.$$

La formule de Stirling donne :

$$\log(30!) = 30{,}5 \cdot \log 30 - 30 \cdot 0{,}4342944819 + 0{,}3990899367$$
$$= 32,\ 422\ 453\ 748\,0 = \log 264^{\text{V}}\ 517\ 093^{\text{IV}} \cdot 10^{24}.$$

La différence, $(265\,252\,859 - 264\,517\,093) \cdot 10^{24} = 735\,766 \cdot 10^{24}$ environ, est un nombre de 30 chiffres ; cette erreur absolue énorme n'est cependant que le 2,8 $^0/_{00}$ du nombre à évaluer.

49. — α) Les exemples ci-dessus mettent en lumière un caractère très curieux de la formule de Stirling : plus n est grand, plus l'erreur commise devient considérable ; *l'erreur absolue croît avec n au-delà de toute limite.* Par contre, *l'erreur relative tend vers zéro, quand* $n \to \infty$. Le tableau suivant le montre (46).

Pour $n =$

5	10	20	30	100

l'erreur absolue commise par l'emploi de (II 33) ou de (II 34) est environ

2	3.10^4	10^{17}	7.10^{30}

tandis que l'erreur relative est à peu près

1,6 $^0/_0$	0,83 $^0/_0$	0,41 $^0/_0$	0,23 $^0/_0$	0,08 $^0/_0$

du nombre à évaluer (47). Comme c'est l'erreur relative qui importe dans la pratique, on peut remplacer $n!$ par le second membre de la formule (II 33).

β) On obtient une approximation plus grande en tenant compte du deuxième terme de la formule (II 32). Prenant les logarithmes népériens des deux membres et désignant par $ln\,x$ le logarithme népérien de x, on trouve (48)

$$ln(n!) = \left(n + \frac{1}{2}\right) . ln(n) - n + \frac{1}{2} . ln(2\pi) + \frac{1}{12n}, \quad \text{(II 36)}$$

où $$\frac{1}{2} . ln(2\pi) = 0{,}918\,93853.$$

(46) Voir Guido Castelnuovo, *Calcolo delle probabilità*, Milan, Rome et Naples, 1919, p. 66. Emanuel Czuber, *Wahrscheinlichkeitsrechnung und ihre Anwendung...*, t. I, 3e éd. Leipzig et Berlin, 1914, p. 27.

(47) On retrouve un caractère analogue dans le comportement de l'écart absolu et de l'écart relatif d'une série d'épreuves répétées ; voir chap. IV, n° 81, δ).

(48) Emile Borel, *Eléments de la théorie des probabilités*, 2e éd. Paris, 1910, p. 44/45.

Enfin, une troisième approximation est fournie par la formule

$$ln(n!) = \left(n + \frac{1}{2}\right).ln(n) - n + \frac{1}{2}\,ln(2\pi) + \frac{1}{12n} - \frac{1}{360n^3} \quad \text{(II 37)}$$

Si l'on ne dispose pas d'une table de logarithmes népériens, on peut remplacer, dans les deux dernières formules, le logarithme népérien $ln(n)$ par le logarithme vulgaire $\log n$, grâce à la relation $ln\,x = \log x\,.\,ln\,10 = \log x.2,\,30258\;50929\;94045\;68401\;7991\ldots.$ Ayant calculé $ln(n!)$ par la formule (II 37), on en déduira le logarithme vulgaire, $\log(n!)$, en multipliant par le module $M = \log e$;

$$\log x = lnx.\ 0{,}434\;294\;481\;903\;251\;827\;651\;129\ldots.$$

L'erreur commise sur $ln(n!)$ par l'emploi de (II 37) est de l'ordre de grandeur de $1 : (360.n^3)$.

γ) En partant de la formule sommatoire d'Euler-MacLaurin donnant le développement connu

$$ln(x!) = \left(x + \frac{1}{2}\right).lnx + \frac{1}{2}\,ln(2\pi) - x + \frac{1}{12x} - \frac{1}{360x^3} + \frac{1}{1260x^5} - \frac{1}{1680x^7} + \frac{1}{1188x^9} - + \ldots\ldots \quad \text{(II 38)}$$

Mr E.-B. Escott ([48 a]) arrive à la formule approchée que voici, bien plus précise que les précédentes :

$$\log(x!) = \left(x + \frac{1}{2}\right).\left\{\frac{12}{7}\log\left(x + \frac{1}{2}\right) - \frac{5}{14}\log\left(x^2 + x + \frac{11}{30}\right) - \log e\right\} + \frac{1}{2}\log(2\pi) \quad \text{(II 39)}$$

Elle subsiste si l'on y remplace partout les logarithmes ordinaires par les logarithmes naturels, $\log x$ par lnx. Elle donne par exemple de 50 !, qui a 65 chiffres, les 11 premiers chiffres exacts, alors que (II 33) en donne seulement les 2 premiers chiffres.

(48 *a*) v. *The American math. Monthly*, vol. 30, 1923, p. 439-441.

CHAPITRE III

LES SÉRIES D'ÉPREUVES RÉPÉTÉES

I. — *La loi empirique du hasard ou loi des grands nombres.*

50. — Pour l'instant, je ne définirai pas le hasard. Ce mot éveille dans l'esprit une idée suffisamment claire pour permettre d'y appliquer le raisonnement. On peut en effet raisonner sans équivoque sur un sujet même vaguement défini, de même qu'on peut étudier le courant électrique sans avoir défini au préalable ce que c'est que l'électricité, ou qu'on peut raisonner juste sur des figures géométriques même vaguement tracées. Le hasard paraît à première vue l'antithèse de toute loi (on a même voulu le définir ainsi), et pourtant il a sa loi. Elle est si nette et si permanente que, mise sous forme mathématique, elle pourrait servir à le définir. On l'appelle *la loi des grands nombres* ou *loi empirique du hasard*. Trois notions fondamentales y interviennent : celle de fréquence relative, celle d'écart ou fluctuation, celle d'absence de régularité. Un exemple imaginé par Joseph Bertrand fait saisir vivement cette loi [49]. S'il pleut un jour entier sur une place non couverte, tous les pavés seront également mouillés. « Il pourrait se faire assurément, lorsque tout alentour la pluie tombe à torrents, qu'un certain pavé restât sec. Aucune goutte n'a pour lui de destination précise, le hasard les disperse, il peut les porter toutes sur les pavés voisins. Personne cependant ne le supposera sérieusement. Telle est la puissance des grands nombres. Le hasard a des caprices, jamais on ne lui vit d'habitudes. Si 1000 gouttes tombent

(49) Joseph Bertrand, *Calcul des probabilités*, Paris, 1889, préface page XIV.

sur 1000 pavés, chaque pavé n'aura pas la sienne; s'il en tombe 1000 millions, chaque pavé recevra son million ou bien peu s'en faudra ».

La loi des grands nombres est un fait expérimental constaté des milliers et des millions de fois. Je n'essaierai donc pas de le démontrer. On ne donne pas la démonstration d'un fait par des raisonnements basés sur une théorie logique ; le fait expérimental s'impose. Par contre, on peut chercher à l'expliquer, à en trouver les causes ou la raison d'être ; c'est à quoi sert le théorème de Jacques Bernoulli. De même qu'on doit voir dans la théorie atomique une explication des phénomènes chimiques, non une démonstration ni une preuve de leur réalité, on doit voir dans le théorème de Jacques Bernoulli, quoiqu'il se démontre comme tous les théorèmes mathématiques, une explication théorique de la loi des grands nombres, mais non une démonstration de sa réalité. Celle-ci, je le répète, est un fait d'expérience (v. n° 54).

51. — L'exemple des pavés également mouillés met en lumière le premier caractère de la loi des grands nombres : la constance de la fréquence relative. Joseph Bertrand en fait aussi ressortir un deuxième caractère distinctif : les fluctuations. Il continue ainsi : « D'un autre côté, il serait fort étrange que les pavés, quoique mouillés également chacun par son million de gouttes, n'eussent pas reçu, dans le cours d'une journée, quelques centaines de gouttes en plus ou en moins. De même, sur quelques millions de coups de dés, quelques points se montreront sans doute un peu plus, d'autres un peu moins souvent. Les rapports sont certains, non les différences ». La formulation précise de ces deux premiers caractères permettra de mettre en lumière un troisième trait distinctif de la loi des grands nombres : l'espèce particulière d'irrégularité qu'elle implique (v. les numéros 59, 60, 191, 204). Commençons par poser quelques définitions.

52. — Soumettons un événement fortuit E à une longue série d'épreuves, répétées r fois de suite dans des conditions autant que possible identiques. Supposons, pour fixer les idées, que E se réalise en tout 55 fois. Ce nombre est dit *le nombre de fréquence* de E

ou sa *fréquence absolue*. Pour juger si E est un événement rare ou non, la connaissance de sa fréquence absolue, 55 dans notre exemple, est insuffisante, il faut connaître sa *fréquence relative*. Notre présomption pour l'arrivée de E sera tout autre en effet, si nous savons que les 55 réalisations de l'événement se sont produites sur 60 épreuves ou sur 110 ou seulement sur 11000. La fréquence relative de E est, dans ces trois hypothèses, respectivement 55 : 60 ou $\frac{11}{12}$; 55 : 110 ou $\frac{1}{2}$; 55 : 11000 ou $\frac{1}{200}$.

Définitions 1 et 2. — Si, dans une série de r épreuves, un événement E qui dépend du hasard s'est réalisé k fois, k est dit son *nombre de fréquence* ou sa *fréquence absolue*. La fraction numérique $k : r$, rapport du nombre de fréquence au nombre total des épreuves, est dite *la fréquence relative* de E.

Corollaires. — 1) La fréquence relative est toujours un nombre compris entre 0 et 1, puisque le nombre des apparitions, k, ne peut pas être supérieur à r, nombre total des épreuves effectuées.

2) La fréquence relative peut exceptionnellement prendre les valeurs limites, savoir la valeur 1 quand l'événement étudié s'est présenté à chaque épreuve, la valeur 0 quand l'événement E ne s'est réalisé dans aucune des épreuves effectuées.

Remarque. — La définition des fréquences, contrairement à celle de la probabilité mathématique, ne présente aucune difficulté théorique ; il suffit de savoir compter.

53. — Pour un même événement, la fréquence varie en général avec le nombre d'épreuves, r. Supposons effectuées, toujours dans les mêmes conditions, une suite de t séries comprenant respectivement $r_1, r_2, r_3, \ldots\ldots, r_t$ épreuves. Soient $k_1, k_2, k_3, \ldots\ldots, k_t$ les nombres de fréquence correspondants de l'événement E. On obtiendra pour sa fréquence relative une suite de t nombres :

$$\frac{k_1}{r_1}, \frac{k_2}{r_2}, \frac{k_3}{r_3}, \ldots\ldots, \frac{k_t}{r_t}, \qquad (a)$$

en général différents les uns des autres ; et même si l'on maintient

constant le nombre r et qu'on effectue pour la deuxième, pour la troisième fois etc., une série de r épreuves, les nombres de fréquence correspondants, k'', k''' etc., seront en général différents, de même que les fréquences relatives de E,

$$\frac{k}{r}, \frac{k'}{r}, \frac{k''}{r}, \frac{k'''}{r}, \qquad (b)$$

L'expérience enseigne toutefois que, pour une classe très étendue de phénomènes du moins, les fréquences relatives dans les suites (a) et (b) oscillent autour d'un certain nombre p qui, chose remarquable, coïncide avec la probabilité mathématique de l'événement étudié. Ce fait expérimental constitue le premier caractère de la loi du hasard. Il y a plus : l'expérience enseigne encore que la fréquence relative se rapproche d'autant mieux de la probabilité mathématique que le nombre d'épreuves, r, est plus grand ; la probabilité mathématique est en un certain sens la limite vers laquelle tend la fréquence relative, lorsque les séries d'épreuves s'allongent indéfiniment (v. les numéros 189, 202, 203).

54. — Cette constatation jette un pont entre les phénomènes du monde matériel et les propositions abstraites du calcul des probabilités. De même qu'aucun raisonnement abstrait ne permet de prévoir que l'eau se dilate en se solidifiant, de même on ne peut pas *démontrer* que l'extraction de 200 000 boules d'une urne $U(1, 1 ; 2)$ dans les conditions du n° 33, définition 8, donnera des boules des deux couleurs dans une proportion sensiblement égale ; de même encore, rien n'oblige *logiquement* à admettre qu'un dé à jouer soigneusement fabriqué présentera la face marquée 1 à peu près 100 fois dans une série de 600 coups, à peu près 1000 fois dans une série de 6000 coups, à peu près r fois dans une série de $6r$ coups, et cela d'autant plus exactement que r est un nombre plus grand. *A priori* et logiquement, on pourrait admettre d'autres possibilités, puisque les combinaisons sont équivraisemblables. C'est l'expérience qui décide entre les diverses possibilités logiques.

Quand on fait une application concrète des théorèmes de la géométrie ou de la mécanique rationnelle, on a recours, implicitement il est vrai, à une proposition du genre de celle-ci : « Les théorèmes

relatifs aux lignes mathématiques sont applicables très approximativement aux fils tendus et aux droites tracées sur un dessin », ou bien : « Les théorèmes sur les corps rigides s'appliquent approximativement aux règles divisées ». Il en va de même dans la théorie des chances. On admet implicitement un axiome disant, en substance, que les théorèmes du calcul des probabilités s'appliquent avec une approximation plus ou moins grande aux phénomènes naturels qui se répètent dans des conditions à peu près identiques. C'est cet axiome qu'on appelle *la loi empirique du hasard* ou loi des grands nombres. On verra plus bas qu'il a servi de point de départ pour arriver à la définition moderne de la probabilité mathématique. On peut l'énoncer comme suit : *Axiome. Dans une série d'épreuves répétées un grand nombre de fois dans les mêmes conditions, chacun des événements fortuits possibles se manifeste avec une fréquence relative sensiblement égale à sa probabilité mathématique. L'approximation augmente, en général assez rapidement, à mesure que les épreuves deviennent plus nombreuses* (v. les numéros 189 et 202).

Remarque. — Dans l'énoncé de cette loi du hasard figurent des locutions imprécises : « sensiblement égale », « augmente, en général ». Une imprécision de ce genre est inévitable toutes les fois qu'on établit une relation entre des concepts abstraits et des phénomènes concrets. Ceci nous amène à une deuxième notion fondamentale dans cette théorie, la notion d'écart.

II. — *Les divers concepts d'écart.*

55. — Supposons que dans une série de 400 épreuves à pile ou face, l'événement E, apparition de face, se soit produit 220 fois et son contraire, apparition de pile, 180 fois. La différence entre la fréquence relative de E et sa probabilité mathématique qui est ici exactement $\frac{1}{2}$ (v. n° 38), s'appelle *l'écart relatif de la série*. Dans notre exemple, l'écart relatif est

$$\frac{220}{400} - \frac{1}{2} = \frac{11}{20} - \frac{10}{20} = \frac{1}{20} = 0{,}05.$$

Autre exemple. On a lancé 12 000 fois un dé à jouer. L'événement E, apparition du point 6, est arrivé 1991 fois ; 1991 $\equiv$ le nombre de fréquence. La fréquence relative de E est donc $\equiv$ 1991 : 12000, et puisque sa probabilité mathématique $p(E) = 1/6$ (v. n° 23, problème 1), l'écart de la série est

$$\frac{1991}{12000} - \frac{1}{6} = \frac{-9}{12000} = -\frac{3}{4000} = -0{,}00075.$$

A toute série d'épreuves répétées dans des circonstances à peu près identiques et portant sur un événement donné E, ou à toute série d'observations exécutées dans les mêmes conditions et portant sur la mesure d'une grandeur, correspond ainsi un écart déterminé ; d'où la

Définition 3. — Un événement fortuit E étant soumis à une série d'épreuves répétées r fois de suite, *l'écart relatif de la série*, L, est la différence entre la fréquence relative de l'événement étudié dans la série en question et sa probabilité mathématique $p(E)$;

$$L \equiv \text{écart relatif} \equiv \frac{k}{r} - p. \qquad \text{(III 1)}$$

L'écart relatif d'une série peut être positif, nul ou négatif. Il joue un rôle prépondérant en stocastique. Il est la différence de deux termes dont le premier est variable d'une série à l'autre, tandis que le deuxième, p, est un nombre fixe calculable à priori. Les changements du nombre de fréquence entraînent les changements de l'écart. *Exemple.* Dans une suite de 12 séries, chacune de 100 coups à pile ou face, l'écart relatif est resté inférieur à 7 %. Que peut-on en conclure sur le nombre de fréquence d'apparition de face ? *Solution* : la fréquence relative est restée entre les limites 0,43 et 0,57 ; donc, dans chacune de ces séries de 100 épreuves, face est apparue au moins 43 et au plus 57 fois.

Corollaires. — 1) L'écart relatif, différence de deux fractions positives proprement dites, est toujours compris entre — 1 et + 1.

2) En utilisant cette notion, on peut énoncer comme suit la loi empirique du hasard. *Axiome* : Dans une série d'épreuves répétées un grand nombre de fois dans les mêmes conditions, chacun des

événements fortuits possibles se réalise avec une fréquence relative sensiblement égale à sa probabilité mathématique. *L'écart relatif d'une série tend stocastiquement vers zéro, quand le nombre d'épreuves augmente indéfiniment.* Pour la notion de limite stocastique, voir les numéros 59, 60, 207.

56. — Nous avons distingué entre la fréquence relative et la fréquence absolue ; cette dernière est le produit de la première par le nombre d'épreuves, r (v. n° 52). Il y a lieu de faire une distinction analogue dans la théorie des écarts et d'introduire, à côté de la notion d'écart relatif, celle d'écart absolu d'une série, par la

Définition 4. — *L'écart absolu* ou *la fluctuation d'une série* de r épreuves faites dans les mêmes conditions est le produit de l'écart relatif par le nombre d'épreuves,

$$l \equiv \text{écart absolu} \equiv \left(\frac{k}{r} - p\right) . r = k - p.r \qquad \text{(III 2)}$$

Si $p.r$ n'est pas un nombre entier, l'écart absolu est la différence entre le nombre de fréquence k et le plus grand nombre entier contenu dans $pr + p$.

Définition 5. — Je désignerai par $En(x)$ le plus grand nombre entier contenu dans x. Exemples : $En(5\ ^1/_2) = 5$; $En(\pi) = 3$; $En(-e) = -2$; $En(\log 7891) = 3$. En utilisant ce symbole, on peut écrire :

$$\text{écart absolu} \equiv l = k - En(pr + p). \qquad \text{(III 3)}$$

Ajoutons ici, par anticipation, que l'écart absolu est la différence entre le nombre de fréquence réellement observé, k, et le nombre de fréquence le plus probable de la série d'épreuves (v. n° 70).

57. — Dans les applications, on est amené à *mesurer* l'écart. Il y a donc lieu d'introduire une unité d'écart. On a posé pour cela les définitions suivantes. Envisageons de nouveau une série de r épreuves exécutées dans les mêmes conditions ; à chaque épreuve, l'événement E a la probabilité mathématique supposée constante $p(E) = p$ et l'événement contraire, F, la probabilité mathémati-

que également constante $p(F) = q$, de se réaliser, en sorte que $p + q = 1$.

Définition 6. — *L'unité d'écart absolu* ou *unité de fluctuation* est la quantité

$$u \equiv \sqrt{2pqr} = \sqrt{2p(1-p)r}. \qquad \text{(III 4)}$$

Il suit de cette définition que l'unité d'écart absolu augmente comme la racine carrée du nombre des épreuves.

On passe de l'unité d'écart absolu à l'unité d'écart relatif comme on passe de la fréquence absolue à la fréquence relative : en prenant le rapport au nombre d'épreuves, c'est-à-dire en divisant par r, nombre total des épreuves exécutées (v. n° 52).

Définition 7. — *L'unité d'écart relatif* est la quantité

$$\frac{u}{r} = \frac{\sqrt{2pqr}}{r} = \sqrt{\frac{2pq}{r}} = \sqrt{\frac{2p(1-p)}{r}} \qquad \text{(III 5)}$$

On l'appelle souvent l'*écart étalon*. Il suit de cette définition que l'unité d'écart relatif, contrairement à l'unité d'écart absolu, est d'autant plus petite que le nombre d'épreuves est plus grand. *L'écart étalon est inversement proportionnel à la racine carrée du nombre des épreuves.* Si l'on fait des séries de cent en cent fois plus longues, l'unité d'écart relatif sera de dix en dix fois plus petite.

Corollaire. — Pour un nombre d'épreuves donné et fixe, r, les deux unités d'écart sont maximales quand $p = q = \frac{1}{2}$. Elles deviennent plus petites et tendent vers zéro, r restant fixe, quand la probabilité mathématique $p \to 0$ ou quand $p \to 1$.

58. — *Définition 8.* — Le nombre qui mesure l'écart absolu l, c'est-à-dire qui indique combien de fois l contient l'unité d'écart absolu, s'appelle **l'écart réduit de la série**. En le désignant par λ, on a donc

$$\lambda \equiv \text{écart réduit} \equiv \frac{l}{u} = \frac{l}{\sqrt{2pqr}} \qquad \text{(III 6)}$$

Exprimé en fonction de l'écart relatif L de la série, l'écart réduit a pour expression

$$\lambda = \frac{L \cdot r}{\sqrt{2pqr}} = L \cdot \sqrt{\frac{r}{2pq}} = L \cdot \sqrt{\frac{r}{2p(1-p)}} \qquad \text{(III 7)}$$

Dans la théorie des écarts, l'écart réduit a une importance capitale. Elle réside dans ce *Théorème. L'écart réduit mesure à la fois l'écart absolu et l'écart relatif d'une série.* Démonstration : Mesurer l'écart relatif L signifie déterminer combien de fois il contient l'unité d'écart relatif. Or, celle-ci est $\equiv u : r$ (v. n° 57, définition 7). Pour trouver combien de fois L contient u/r, il suffit de faire le rapport de L à u/r, ce qui donne

$$L : \frac{u}{r} = L \cdot \frac{r}{u} = \frac{L\ r}{u} = \frac{l}{u}$$

qui n'est pas autre chose que l'écart réduit λ, en vertu de (III 6).

Exemple 1. — Dans une série de 900 épreuves à pile ou face, l'événement E, apparition de face, s'est produit 420 fois.

1) Le nombre de fréquence est donc $k = 420$, c'est la fréquence absolue de E.

2) La fréquence relative, puisque $r = 900$, est

$$\frac{k}{r} = \frac{420}{900} = \frac{7}{15} = 0{,}46666\ldots$$

3) La probabilité mathématique étant $p(E) = \frac{1}{2}$ (v. n° 38), on a, pour l'écart relatif de la série,

$$L = \frac{k}{r} - p = \frac{7}{15} - \frac{1}{2} = \frac{-1}{30} = -\ 0{,}03333\ldots$$

4) L'écart absolu ou la fluctuation de la série est ici

$$l \equiv k - pr = 420 - \frac{1}{2} \cdot 900 = 420 - 450 = -\ 30.$$

C'est la différence entre le nombre de fréquence réellement observé, 420, et la fréquence la plus probable qui est ici 450 (v. n° 70).

5) L'unité d'écart absolu étant

$$u = \sqrt{2 \cdot 1/2 \cdot 1/2 \cdot 900} = \sqrt{450} = 15 \cdot \sqrt{2},$$ on a

6) L'unité d'écart relatif, ou l'écart étalon,

$$\frac{u}{r} = \frac{15 \cdot \sqrt{2}}{900} = \frac{1{,}41421356}{60} = 0{,}023570$$

7) Enfin, *l'écart réduit* λ devient

$$\lambda = l : u = -30 : (15\sqrt{2}) = -\sqrt{2} = -1{,}41421356...$$

Exemple 2. — Dans une série de 1000 coups de dé, le six fut amené 215 fois. On a donc 1) le nombre de fréquence $k = 215$; c'est la fréquence absolue de l'événement étudié $E \equiv$ apparition du point 6.

2) La fréquence relative de E est $\frac{k}{r} = \frac{215}{1000} = \frac{43}{200} = 0{,}215$.

3) La probabilité mathématique étant $p = \frac{1}{6}$ (v. n° 23, problème 1), on a pour l'écart relatif de cette série de $r = 1000$ coups,

$$L = \frac{k}{r} - p = \frac{43}{200} - \frac{1}{6} = \frac{29}{600} = 0{,}048333...$$

4) Pour l'écart absolu, l, de cette série, puisque $p \cdot r = 1000 \cdot \frac{1}{6}$ n'est pas un nombre entier, il faut prendre le plus grand entier contenu dans $pr + p = 1001 : 6 = 166 + 5/6$; c'est 166 ; donc, en vertu de (III 3),

$l = 215 - 166 = 49$. C'est de nouveau la différence entre le nombre de fréquence réellement observé, 215, et celui que la théorie fait prévoir, 166.

5) L'unité d'écart absolu est

$$u = \sqrt{2 \cdot \frac{1}{6} \cdot \frac{5}{6} \cdot 1000} = \frac{1}{6} \cdot 10 \cdot \sqrt{100} = \frac{50}{3}.$$

6) L'unité d'écart relatif est

$$\frac{u}{r} = \frac{50}{3} : 1000 = \frac{1}{60} = 0{,}016666...$$

7) Enfin *l'écart réduit*, λ, devient

$$\lambda = \frac{l}{u} = 49 : 50\sqrt{3} = \frac{49.3.2}{50.2} = 2{,}94.$$

Remarque. — En comparant les deux écarts réduits, 1,414... et 2,94 on peut dire, en vertu du théorème ci-dessus, que la seconde série est plus de 2 fois plus irrégulière que la première.

La limite stocastique.

59. — Il reste à mettre en lumière le troisième caractère de la loi empirique du hasard (v. n° 51). Dans les sciences exactes, on peut, sinon connaître la grandeur de l'erreur commise, du moins en fixer une limite supérieure. C'est le cas par exemple lorsqu'on fait une multiplication abrégée ou qu'on emploie une table de logarithmes à 5 décimales ou que, dans une série infinie, on ne prend que les n premiers termes. Ainsi, dans la formule de Stirling, on peut assigner une limite supérieure de l'erreur relative (v. n° 49). Or, rien de semblable n'a lieu dans la théorie des probabilités. Même si l'on répète l'épreuve : le lancement de la pièce de monnaie ou le tirage d'une boule de l'urne $U(a, a ; s = 2a)$ avec les conditions du n° 33, définition 8, dix mille fois, un million de fois,....., 10^n fois, les fréquences relatives de face ou de boule blanche,

$$\frac{k_4}{10\,000}, \frac{k_5}{100\,000}, \frac{k_6}{1\,000\,000}, \ldots\ldots, \frac{k_n}{10^n},$$

oscilleront autour du nombre 1/2 ; mais il est impossible de prédire *à coup sûr* une limite que l'écart ne dépassera jamais ; en d'autres termes, on ne peut pas fixer un nombre N de coups tel que la fréquence relative correspondante soit comprise *sûrement* entre 0,47 et 0,53 par exemple. En théorie, il est extrêmement invraisemblable et pratiquement tout-à-fait exclu que, sur un million d'épreuves, la fréquence relative tombe en dehors de cet intervalle. Déjà pour une série de 10 000 épreuves à pile ou face, la probabilité mathématique d'un écart supérieur à 3 % est moindre que 1 : 50 000 000 ; mais, si follement invraisemblable qu'il soit, un tel écart n'est pas

logiquement absurde ni mathématiquement exclu (50) et cela quelque grand que soit le nombre d'épreuves N. Pour exprimer ce fait sous une forme concise, je dirai que l'écart relatif *tend stocastiquement vers zéro* (v. les numéros 60, 207).

60. — En analyse mathématique, une suite illimitée de grandeurs $x_1, x_2, \ldots\ldots, x_n, \ldots\ldots$ étant donnée, le symbole

$$x_n \to a$$

exprime simultanément les deux faits suivants : 1° Quand $n \to \infty$, la différence $(x_n - a)$, prise en valeur absolue, finit par devenir plus petite que tout nombre positif ε choisi d'avance aussi petit que l'on veut ; 2° la valeur absolue de cette différence $(x_n - a)$ reste inférieure à ε pour *tous* les indices n au delà d'une certaine limite, dès que $n > N$, où N est un nombre positif qui peut dépendre de ε. Rien de semblable dans le calcul des probabilités ; la locution : « Cette grandeur variable tend vers la probabilité mathématique » doit y être prise dans le sens moins rigoureux qu'elle a dans les sciences expérimentales, lorsqu'on dit par exemple que le mélange de deux liquides tend à devenir homogène avec le temps ou que la température de deux corps voisins tend à s'égaliser. C'est ce caractère imprécis que doit marquer l'adverbe *stocastiquement*.

On peut donc énoncer la loi empirique du hasard comme suit. *Axiome. Dans une série d'épreuves répétées dans des conditions identiques, chacun des événements fortuits possibles, E, se réalise avec une fréquence relative qui tend stocastiquement vers sa probabilité mathématique p(E), quand le nombre d'épreuves croît indéfiniment.*

Remarque. — Dans le calcul des probabilités classique, cet axiome ne sert pas de base à l'édifice logique de la théorie des chances, mais seulement de lien entre la théorie abstraite et ses applications au monde matériel. La théorie des chances apprend à déterminer des probabilités mathématiques, concepts abstraits,

(50) E. Carvallo. *Le calcul des probabilités et ses applications*, Paris, 1912, p. 10-11.

tandis que la pratique demande de prévoir des fréquences, réalisations concrètes d'événements physiques, donc tout autre chose que des abstractions. Cependant, bien que les affirmations du théorème de Jacques Bernoulli se rapportent à des probabilités mathématiques et non à des nombres de fréquence, la théorie permet, dans certains cas très importants, de prévoir les événements avec une certitude qui n'est jamais dépassée dans les connaissances humaines (v. n° 82).

Ces préliminaires posés, nous allons aborder le théorème de Jacques Bernoulli.

III. — *Les épreuves répétées et la formule du binôme de Newton.*

61. — Il existe une analogie remarquable entre les formules qui donnent les probabilités mathématiques se rapportant au concours de plusieurs événements et celles qui donnent le produit de plusieurs binômes ou polynômes. Cette analogie, remarquée déjà par Jacques Bernoulli, développée par Abram de Moivre et surtout par Laplace, permet d'obtenir immédiatement les probabilités des différentes combinaisons.

Pour commencer par le cas le plus simple, envisageons : 1) l'événement E de probabilité mathématique p et l'événement contraire, F, de probabilité mathématique $q = 1 - p$; 2) l'événement E' dont la probabilité mathématique soit $p(E') \equiv p'$ et l'événement contraire, F', de probabilité mathématique $p(F') \equiv q' = 1 - p'$. En vertu du théorème des probabilités composées (v. les numéros 33 et 34), le produit

$pp' = p(EE')$ = prob. math. que E et E' arrivent tous les deux ;
$pq' = p(EF')$ = » » » E, mais non E' se réalise ;
$qp' = p(FE')$ = » » » E', mais non E arrive ;
$qq' = p(FF')$ = » » » ni E ni E' ne se réalisent.

Or, en vertu des règles de l'algèbre élémentaire,

$$(p + q)(p' + q') = pp' + pq' + qp' + qq'. \qquad \text{(III 8)}$$

Les divers termes de ce développement ont donc une signification bien déterminée dans le calcul des probabilités ; c'est en cela que réside l'analogie en question.

Remarque. — D'après le théorème d'addition (v. n° 32), on voit que

$pp' + pq' + p'q$ = probabilité mathématique que *l'un au moins* des deux événements E et E' se réalise ;

$pq' + p'q$ = probabilité mathématique que *l'un seulement* des deux événements arrive ;

$pq' + p'q + qq'$ = probabilité mathématique que *l'un au plus* des deux événements arrive.

62. — Faisons en particulier $E' = E$; il s'agit alors de la répétition d'un même événement E. L'hypothèse que les circonstances produisant E restent toujours les mêmes entraîne $p' = p$, $q' = q$, et la formule (III 8) devient

$$(p + q)^2 = p^2 + 2pq + q^2.$$

De nouveau, chaque terme de ce développement a une signification précise dans la théorie des chances, savoir : Que l'événement étudié arrive, dans une série de $r = 2$ épreuves,

2 fois de suite, la probab. math est p^2,
1 » seulement, » » » $2pq$,
0 » » » » q^2. En outre,

$p^2 + 2pq$ = » » que E arrive au moins une fois,
$2pq$ = » » » E » exactement » ;
$q^2 + 2pq$ = » » » E » au plus » .

Supposons qu'on exécute une série de 3 épreuves, dans les mêmes conditions. Par le même raisonnement, on démontrerait que dans le développement

$$(p + q)^3 = p^3 + 3p^2q + 3pq^2 + q^3, \qquad \text{(III 9)}$$

chaque terme a une signification précise au point de vue de la

théorie des chances, savoir : Que dans une série de $r = 3$ épreuves, l'événement E se réalise

3 fois de suite, la probab. math. est p^3,
2 » exactement, » » » » $3p^2q$,
1 » » , » » » » $3pq^2$,
0 » , » » » » q^3.

Comme l'une de ces quatre possibilités se réalise sûrement, la somme des probabilités doit égaler 1 ; et c'est en effet le cas, puisque $(p + q)^3 = 1^3 = 1$.

Remarque 1. — Dans tous ces développements, l'ordre de succession des arrivées de E et de F n'est pas prescrit d'avance et par conséquent quelconque. Voir à ce sujet au numéro 38 la solution du problème 14.

Remarque 2. — En vertu du théorème d'addition (v. n° 32), la probabilité mathématique que, dans une série de $r = 3$ épreuves, l'événement E se réalise

1 fois au moins, est $= p^3 + 3p^2q + 3pq^2 = 1 - q^3$ (v. n° 35, probl. 8),
2 » » » , » $= p^3 + 3p^2q$,
1 » au plus, » $= q^3 + 3pq^2$,
2 » » » , » $= q^3 + 3pq^2 + 3p^2q$.

63. — Soumettons à une série de r épreuves, répétées dans les mêmes conditions, un événement fortuit E dont la probabilité mathématique soit constamment $= p$ et par conséquent la probabilité de non-réalisation constamment $1 - p \equiv q = p(F)$. Avec notre notation symbolique (v. n° 31, définition 5), on peut écrire, en convenant que l'ordre de succession des événements E et F ne soit pas prescrit d'avance, pour une série de $r = 3$ épreuves :

$$(p + q)^3 = p(E^3) + p(E^2F) + p(EF^2) + p(F^3) ; \quad \text{(III 9')}$$

pour une série de $r = 4$ épreuves, on a :

$$\left.\begin{aligned}(p+q)^4 &= p^4 + 4p^3q + 6p^2q^2 + 4pq^3 + q^4 \\ &= p(E^4) + p(E^3F) + p(E^2F^2) + p(EF^3) + p(F^4),\end{aligned}\right\} \text{(III 10)}$$

c'est-à-dire que chaque terme du développement de $(p+q)^4$ donne précisément la valeur de la probabilité mathématique de l'une ou de l'autre des 5 combinaisons possibles, dans l'arrivée ou la non-arrivée de E en 4 épreuves. De plus, en vertu du théorème d'addition (v. n° 32), la probabilité mathématique pour que, dans une série de 4 épreuves, E se réalise

1 fois au moins, est $= p^4 + 4p^3q + 6p^2q^2 + 4pq^3 = 1 - q^4$ (n° 35, pr. 8),
2 » » » , » $= p^4 + 4p^3q + 6p^2q^2$,
3 » » » , » $= p^4 + 4p^3q$,
1 » au plus, » $= q^4 + 4pq^3$,
2 » » » , » $= q^4 + 4pq^3 + 6p^2q^2$,
3 » » » , » $= q^4 + 4pq^3 + 6p^2q^2 + 4p^3q$.

On verrait de même que chaque terme du développement de $(p+q)^5$ est égal à la probabilité mathématique de l'une des 6 combinaisons possibles dans l'arrivée ou la non-arrivée de E en une série de $r = 5$ épreuves ; et ainsi de suite pour $r = 6, 7, \ldots, n$. Par le procédé dit de l'induction complète, c'est-à-dire par la conclusion de n à $n+1$, on démontre que cette loi est générale. On peut l'énoncer comme suit.

Théorème 1. — Soient p et $q \equiv 1 - p$ les probabilités mathématiques de deux événements contraires E et F. La valeur de l'expression

$$\frac{(k+n)!}{k!\,.\,n!}\,p^k q^n \qquad \text{(III 11)}$$

est égale à la probabilité mathématique que le résultat de $r \equiv k+n$ épreuves identiques consistera dans l'arrivée k fois répétée de E et n fois répétée de F, dans un ordre de succession quelconque et non fixé d'avance.

Faisons parcourir à k la série des nombres 0, 1, 2, 3, ..., r ; les combinaisons correspondantes sont exclusives l'une de l'autre ; on peut donc appliquer le théorème d'addition et dire :

Théorème 2. — Si l'on développe par la formule de Newton

$$\left.\begin{aligned}(p+q)^r = p^r + \frac{r!}{(r-1)!\,1!}\,p^{r-1}q^1 + \frac{r!}{(r-2)!\,2!}\,p^{r-2}q^2 \\ + \frac{r!}{(r-3)!\,3!}\,p^{r-3}q^3 + \ldots\ldots + \frac{r!}{(r-n)!\,n!}\,p^{r-n}q^n + \ldots\ldots + q^r\end{aligned}\right\} \qquad \text{(III 12)}$$

chaque terme de ce développement représente la probabilité mathématique d'une certaine combinaison, savoir : le terme général (III 11) donne la probabilité mathématique que l'événement étudié E, de probabilité mathématique p, se réalise $r - n = k$ fois et l'événement contraire n fois, dans un ordre de succession quelconque non fixé d'avance. Ceci suppose que l'on exécute une série de $r = k + n$ épreuves dans des conditions identiques, en particulier que les probabilités p et $q \equiv 1 - p$ ne varient pas d'une épreuve à l'autre.

IV. — *La première partie du théorème de Jacques Bernoulli.*

64. — Reprenons le schéma de l'urne avec les conventions du numéro 33, définition 8, l'événement E étant ≡ sortie de boule blanche et l'événement contraire, F, sortie de boule noire. Si l'on fait abstraction de l'ordre de succession, une série de r épreuves donne lieu à $r + 1$ combinaisons exclusives l'une de l'autre, savoir :

1) E se produit r fois et F arrive 0 fois
2) E » » $r - 1$ » » F » 1 »
3) E » » $r - 2$ » » F » 2 »

. . . . . .

. . . . . .

r) E » » 1 » » F » $r - 1$ »
$r+1$) E » » 0 » » F » r » de suite.

Deux questions fondamentales se posent. *Question A* : parmi ces $r + 1$ possibilités, quelle est la combinaison la plus probable ? et quelle en est la probabilité mathématique ? *Question B* : quelle est la probabilité mathématique P pour que l'écart de la série ou la fluctuation ne dépassent pas certaines limites données d'avance ? En d'autres termes, quelle est la probabilité mathématique que le résultat de ces r épreuves soit la réalisation ou bien de la combinaison la plus probable ou bien d'une combinaison différente, mais dont l'écart reste inférieur à une limite fixée d'avance ? Le

théorème de Jacques Bernoulli permet de répondre à ces deux questions fondamentales ; et voici par quelle méthode. Pour simplifier, nous supposerons le problème ramené au schéma de l'urne.

65. — On fait une série de r épreuves identiques : on tire au hasard, r fois de suite, une boule d'une urne $U(a, b ; s)$ avec les conventions du numéro 33, définition 8.

Question A : Quel est le résultat le plus probable ?

Méthode. — En vertu du théorème ci-dessus (n° 63, théorème 2), la question posée revient à cette autre : Quel est, dans le développement de $(p + q)^r$, le terme de valeur maximale ? Les exposants de p et de q dans ce terme maximal sont les nombres cherchés ; ils indiquent les nombres de fréquence des événements E et F dans la combinaison la plus probable ; et sa probabilité mathématique est donnée par la valeur du terme maximal lui-même.

Question B. Quelle est la probabilité mathématique P_{-l}^{+l} pour que l'écart de la série ne dépasse pas une limite fixée d'avance, en plus ou en moins ?

Méthode. — En vertu du même théorème 2 (v. n° 63), il faut :

1) Déterminer le terme maximal du développement de $(p + q)^r$ par la formule de Newton.

2) Déterminer tous les termes qui, de part et d'autre du terme maximal, ont un écart inférieur ou au plus égal à la limite fixée d'avance.

3) Additionner ces termes-là. La somme ainsi obtenue est la probabilité mathématique cherchée P_{-l}^{+l} (Voir les numéros 80 et suivants).

Commençons par traiter un cas particulier.

66. — *Problème 1.* — D'une urne U (20, 10 ; 30), on fait une série de $r = 6$ tirages, avec les conventions du numéro 33, définition 8. On demande : α) Quel est le résultat le plus probable de ces 6 tirages ? β) Avec quelle probabilité mathématique peut-on s'y attendre ?

Solution. — On a ici $p = \frac{20}{30} = \frac{2}{3}$, $q = \frac{10}{30} = \frac{1}{3}$, $r = 6$.

On doit donc développer

$$\left(\frac{2}{3}+\frac{1}{3}\right)^6 = \left(\frac{2}{3}\right)^6 + 6.\left(\frac{2}{3}\right)^5.\frac{1}{3} + 15.\left(\frac{2}{3}\right)^4.\left(\frac{1}{3}\right)^2 + 20.\left(\frac{2}{3}\right)^3.\left(\frac{1}{3}\right)^3$$
$$+ 15.\left(\frac{2}{3}\right)^2.\left(\frac{1}{3}\right)^4 + 6.\left(\frac{2}{3}\right).\left(\frac{1}{3}\right)^5 + \left(\frac{1}{3}\right)^6$$
$$= \frac{64}{729} + \frac{192}{729} + \frac{240}{729} + \frac{160}{729} + \frac{60}{729} + \frac{12}{729} + \frac{1}{729}.$$

On voit que le terme maximal est $\frac{240}{729} = 15 \,.\, (2/3)^4 \,.\, (1/3)^2$; il correspond aux exposants 4 et 2 ; d'où cette *Réponse* : α) Le résultat le plus probable est 4 fois apparition de boule blanche et 2 fois de boule noire. β) La probabilité mathématique de cette combinaison est 240 : 729 = 0,329 213 992 ..., soit un peu moins d'un tiers.

Problème 2. — De la même urne $U(20, 10 ; 30)$ et dans les mêmes conditions, on fait une série de $r = 12$ tirages. α) Quel est le résultat le plus probable ? β) Quelle en est la probabilité mathématique ?

Solution. — Comme on a $p = 2/3$, $q = 1/3$, $r = 12$, il faut développer

$$\left(\frac{2}{3}+\frac{1}{3}\right)^{12} = \left(\frac{2}{3}\right)^{12} + 12.\left(\frac{2}{3}\right)^{11}.\left(\frac{1}{3}\right) + \frac{12.11}{1.2}.\left(\frac{2}{3}\right)^{10}.\left(\frac{1}{3}\right)^2$$
$$+ \frac{12.11.10}{1.2.3}.\left(\frac{2}{3}\right)^9.\left(\frac{1}{3}\right)^3 + \binom{12}{4}.\left(\frac{2}{3}\right)^8.\left(\frac{1}{3}\right)^4$$
$$+ \ldots\ldots\ldots\ldots$$
$$= \frac{4096}{D} + \frac{24576}{D} + \frac{67584}{D} + \frac{112640}{D} + \frac{126720}{D}$$
$$+ \frac{101376}{D} + \frac{59136}{D} + \frac{25344}{D} + \frac{7920}{D} + \frac{1760}{D}$$
$$+ \frac{264}{D} + \frac{24}{D} + \frac{1}{D}$$

où le dénominateur commun est $D = 3^{12} = 531\,441$. On voit que

le terme maximal est 126 720 : D ; comme c'est le 5ième terme du développement, il correspond aux exposants 8 et 4, ce qui donne cette *Réponse :* α) Le résultat le plus probable de ces 12 tirages est 8 fois boule blanche et 4 fois boule noire. β) La probabilité mathématique de cette combinaison est

$$126\,720 : 531\,441 = 0{,}238\,446\,036 \ldots,$$

soit un peu moins que 6 : 25 ou 24 %.

67. — Avant de passer au cas général, traitons pour ces mêmes exemples la question B du numéro 65.

Problème 3. — De l'urne U (20, 10; 30), on fait une série de $r = 6$ tirages, avec les conventions du numéro 33, définition 8. Quelle est la probabilité mathématique P_{-1}^{+1} que l'écart absolu ne dépasse pas ± 1 ou, ce qui revient au même, que l'écart relatif ne dépasse pas $\pm \frac{1}{6} = \pm 0{,}1666 \ldots = \pm 16\,{}^2/_3\,\%$? *Solution.* Voir le problème 1 du n° 66. La combinaison la plus probable étant $E^4F^2 \equiv 4$ blanches + 2 noires, un écart absolu de ± 1 signifie qu'on sort soit 3 blanches + 3 noires $\equiv E^3F^3$, soit 5 blanches + 1 noire $\equiv E^5F^1$. Comme ces modalités sont exclusives l'une de l'autre, on peut appliquer le théorème d'addition (v. n° 32). Il donne :

$$P_{-1}^{+1} = p(E^5F^1) + p(E^4F^2) + p(E^3F^3)$$
$$= \frac{192}{729} + \frac{240}{729} + \frac{160}{729} = \frac{592}{729} = 0{,}812\,071\,330 \ldots \backsim 9/11.$$

Réponse. — La probabilité cherchée est $P = 0{,}812 \ldots$, soit presque 9/11.

Problème 4. — De la même urne U (20, 10; 30) et dans les mêmes conditions, on effectue $r = 12$ tirages. Quelle est maintenant la probabilité mathématique P que l'écart relatif ne dépasse pas 16 $^2/_3$ % ? *Solution.* Voir le problème 2 du n° 66. Un écart relatif $L = 16\,{}^2/_3\,\% = \frac{1}{6}$ correspond ici à un écart absolu $l = L.r = \frac{1}{6} . 12 = 2$ (v. n° 55 et n° 56). Cela signifie que boule

blanche sortira 8 ± 2 fois, c'est-à-dire de 6 à 10 fois, puisque la combinaison la plus probable contient 8 boules blanches. Les combinaisons pour lesquelles l'écart reste dans les limites prescrites sont donc :

10 blanches + 2 noires, soit $E^{10}F^2$;
9 blanches + 3 noires, soit E^9F^3 ;
8 blanches + 4 noires, soit E^8F^4 ;
7 blanches + 5 noires, soit E^7F^5 ;
6 blanches + 6 noires, soit E^6F^6.

Dès lors, la probabilité mathématique cherchée est donnée par

$$P_{-2}^{+2} = p(E^6F^6) + p(E^7F^5) + \ldots\ldots\ldots + p(E^{10}F^2)$$
$$= \frac{67584 + 112640 + 126720 + 101376 + 59136}{531441}$$
$$= \frac{467456}{531441} = 0{,}87960093\ldots, \text{ soit presque } \frac{22}{25}.$$

68. — Les deux exemples précédents, problèmes 1 à 4, font entrevoir les conclusions suivantes, démontrées plus bas.

a) La probabilité mathématique de la combinaison la plus probable diminue quand le nombre d'épreuves, r, augmente (v. n° 74, I).

b) La probabilité mathématique P que l'écart relatif d'une série reste compris entre des limites fixées d'avance, par exemple qu'il ne dépasse pas c %, augmente avec le nombre d'épreuves r, lorsqu'on maintient ces limites constantes (v. n° 81, γ).

c) Lorsque le nombre d'épreuves r est très grand et les limites de l'écart un peu considérables, les calculs deviennent rebutants par leur longueur et même humainement impossibles, à cause du temps qu'il faudrait y employer. On est alors obligé d'avoir recours à des procédés abréviatifs, bien que théoriquement la question soit parfaitement résolue. Les principaux auxiliaires sont les logarithmes et la formule de Stirling (v. n° 47 et n° 48).

d) Le développement de $(p + q)^r$ donne en général une série *asymétrique*, de sorte que le terme maximal n'est pas le terme équidistant des extrêmes. Si $p > \frac{1}{2}$, le terme maximal est plus rapproché de p^r, premier terme du développement; si $p < \frac{1}{2}$ et par

conséquent $q > \frac{1}{2}$, le terme maximal est plus rapproché de q^r, dernier terme du développement. La série n'est symétrique que dans le cas où $p = q = \frac{1}{2}$; et le terme maximal est alors le terme central si r est un nombre pair; si r est impair et $p = q = \frac{1}{2}$, il y a deux termes maximaux égaux entre eux.

e) L'asymétrie du développement de $(p + q)^r$ peut être si prononcée que le tout premier terme soit déjà le terme maximal. Cela se produit quand le nombre d'épreuves est inférieur à une certaine limite. En étudiant le rapport de deux termes consécutifs, on démontre en effet ce qui suit.

Théorèmes. — 1) Si le nombre d'épreuves $r < \frac{1}{p} - 1$, le terme maximal coïncide avec le dernier terme du développement de $(p + q)^r$; c'est donc q^r, et le résultat le plus probable de la série d'épreuves est la réalisation répétée r fois de suite de l'événement contraire F (v. n° 72, problème 8).

2) Si $r = \frac{1}{p} - 1$, les deux derniers termes du dit développement sont égaux entre eux et sont en même temps les deux termes maximaux, $q^r = r.q^{r-1}.p$ (voir un exemple au numéro 72, problème 8, β).

3) Si le nombre d'épreuves $r < \frac{1}{q} - 1$, le terme maximal est p^r, premier terme du développement en question; et le résultat le plus probable de la série est alors l'arrivée répétée r fois de suite de l'événement étudié E.

4) Si le nombre d'épreuves $r = \frac{1}{q} - 1$, les deux premiers termes sont égaux entre eux et sont en même temps les termes maximaux, $p^r = r.q.p^{r-1}$.

Ces cas exceptionnels que Markoff a signalés le premier ([51]) sont exclus dès que le nombre d'épreuves est suffisamment grand.

(51) A. A. Markoff, traduit en allemand par H. Liebmann, *Wahrscheinlichkeitsrechnung*, 1912, p. 28.

Le résultat le plus probable d'une série d'épreuves identiques

69. — On a vu ci-dessus (v. n° 65) que déterminer le résultat le plus probable d'une série d'épreuves identiques dont chacune réalise soit l'événement E, soit son contraire F, revient à déterminer le terme maximal dans le développement de $(p+q)^r$, où p et $q \equiv 1-p$ sont les probabilités mathématiques de E et de F et r le nombre d'épreuves.

Supposons que le terme maximal, que je désignerai dans la suite par T_o, soit

$$T_o \equiv \frac{r!}{k!.n!} . p^k . q^n \qquad \text{(III 13)}$$

La combinaison la plus probable de notre série est alors celle qui amène, dans un ordre de succession quelconque et non prescrit d'avance, k fois l'événement E et n fois l'événement F, avec

$$k + n = r.$$

α) Le terme T_o donné par (III 13), étant maximal par hypothèse, est plus grand que le terme qui le suit, d'où l'inégalité

$$\frac{r!}{k!.n!} . p^k . q^n > \frac{r!}{(k-1)!.(n+1)!} . p^{k-1} . q^{n+1}$$

Après multiplication par $k!.(n+1)!$ et division par $r!.p^{k-1}.q^n$, il vient

$$(n+1)p > kq \qquad \text{(III 14')}$$

En observant que $q = 1-p$ et $n = r-k$, on en tire

$$k(1-p) < p(r-k+1), \text{ d'où}$$

$$k < pr + p. \qquad \text{(III 14)}$$

β) Le terme T_o, étant maximal par hypothèse, est aussi plus grand que le terme qui le précède, d'où l'inégalité

$$\frac{r!}{k!.n!} . p^k . q^n > \frac{r!}{(k+1)!.(n-1)!} . p^{k+1} . q^{n-1}$$

Après multiplication par $(k+1)!\,n!$ et division par $r!.p^k.q^{n-1}$,

on en déduit

$$(k+1)q > np \qquad \text{(III 14'')}$$

On en tire, puisque $q = 1 - p$ et $n = r - k$,

$$(k+1)(1-p) > p(r-k), \text{ d'où}$$

$$k > pr - q. \qquad \text{(III 15)}$$

Des inégalités (III 14) et (III 15), il suit que k, nombre de fréquence de l'événement étudié, doit être compris entre $pr + p$ et $pr - q$, donc

$$pr - q < k < pr + p. \qquad \text{(III 16)}$$

70. — La différence des limites obtenues pour k est exactement 1, car $(pr + p) - (pr - q) = p + q = 1$. Dès lors, deux cas sont possibles.

1° Les nombres $pr + p$ et $pr - q$ ne sont pas entiers. Il existe alors un seul nombre entier compris entre eux ; donc k, nombre de fréquence le plus probable de E, est déterminé univoquement ; on voit par (III 16) que c'est le plus grand nombre entier contenu dans $pr + p$; la notation du n° 56, définition 5, permet d'écrire

$$k = En(pr + p) = En(p.[r+1]) \qquad \text{(III 17)}$$

En particulier, si $p.r$ est un nombre entier, on a

$$k = pr. \qquad \text{(III 18)}$$

Dans tous ces cas, le développement de $(p+q)^r$ renferme un seul terme maximal. Exemples : $\left(\frac{2}{3} + \frac{1}{3}\right)^6$, v. n° 66 ; ou $\left(\frac{1}{10} + \frac{9}{10}\right)^8$, v. n° 72, problème 8, α) ou γ).

2° Les nombres $pr + p$ et $pr - q$ sont tous les deux entiers. Dans ce cas, le nombre k est déterminé bivoquement, et le développement de $(p+q)^r$ renferme deux termes maximaux égaux entre eux. Exemples : $\left(\frac{2}{3} + \frac{1}{3}\right)^5$, ou $\left(\frac{1}{10} + \frac{9}{10}\right)^9$, v. n° 72, problème 7 et problème 8, β).

On démontrerait de la même manière, en partant des inégalités (III 14') et (III 14''), que n, nombre de fréquence le plus probable

de l'événement F, doit être compris entre les limites $qr + q$ et $qr - p$, de sorte que

$$qr - p < n < qr + q \qquad \text{(III 19)}$$

Si ces deux limites, dont la différence est exactement 1, ne sont pas des nombres entiers, n est déterminé univoquement. C'est le plus grand nombre entier contenu dans $qr + q$, de sorte que

$$n = En(qr + q). \qquad \text{(III 20)}$$

En particulier, si qr est un nombre entier, on a

$$n = qr. \qquad \text{(III 21)}$$

Si les dites limites $qr + q$ et $qr - p$ sont toutes deux des nombres entiers, n est déterminé bivoquement, comme $k = r - n$.

Remarque. — On peut toujours écrire

$$k = pr \pm \alpha, \quad n = qr \pm \alpha' \qquad \text{(III 22)}$$

où α et α' représentent des fractions inférieures à 1/2. Quand le nombre d'épreuves, r, est très grand, on ne commet qu'une erreur négligeable en prenant toujours, même si pr et qr ne sont pas entiers,

$$k = pr, \quad n = qr \qquad \text{(III 23)}$$

comme nombres de fréquence les plus probables de l'événement E et de son contraire F.

71. — En divisant les deux membres de (III 16) et de (III 19) par r, on introduit la fréquence relative (v. n° 52). Il vient

$$p - \frac{q}{r} < \frac{k}{r} < p + \frac{p}{r}, \qquad \text{(III 24)}$$

ce qui signifie que la fréquence relative, $k : r$, de l'événement étudié diffère de sa probabilité mathématique, p, d'une fraction plus petite que $1 : r$. On démontrerait de même, en partant de (III 19), que la différence entre la fréquence relative de l'événe-

ment F et sa probabilité mathématique q est moindre que 1 : r, en formule

$$\frac{n}{r} = q \pm \frac{\alpha'}{r}, \text{ avec } 0 \leqslant \alpha' \leqslant \frac{1}{2};$$

de même

$$\frac{k}{r} = p \pm \frac{\alpha}{r}, \text{ avec } 0 \leqslant \alpha \leqslant \frac{1}{2}.$$

Faisons le rapport des fréquences relatives $\frac{k}{r}$ et $\frac{n}{r}$. Si pr est un nombre entier, ce rapport devient en vertu de (III 23)

$$\frac{k}{n} = \frac{p}{q}. \qquad \text{(III 25)}$$

Si pr n'est pas un nombre entier, le rapport des nombres de fréquence, $k : n$, diffère du rapport des probabilités mathématiques, $p : q$, d'une quantité ε qui $\to 0$ quand $r \to \infty$. On a dès lors la proposition suivante qui constitue la première partie du théorème de Jacques Bernoulli (pour la deuxième partie, voir le numéro 80).

Théorème de Jacques Bernoulli (première partie). — Considérons une série de r épreuves identiques dont chacune produise soit l'événement aléatoire E, soit l'événement contraire F. Supposons que leurs probabilités mathématiques,

$$p(E) \equiv p \text{ et } p(F) \equiv q = 1 - p,$$

restent constantes ; soient enfin k et n les nombres de fréquence de E et de F respectivement. *Le résultat le plus probable de cette série d'épreuves est celui où le rapport des nombres de fréquence*, k : n, *est égal au rapport des probabilités mathématiques*, p : q, *ou en diffère le moins possible*, savoir :

$$k = En(p.r + p), \quad n = En(q.r + q). \qquad \text{(III 26)}$$

On peut énoncer ce résultat plus simplement en introduisant la fréquence relative et formulant comme suit la première partie du théorème de Jacques Bernoulli.

Corollaire. — *Le résultat le plus probable* d'une série de r épreuves dont chacune amène soit l'événement aléatoire E de pro-

babilité mathématique constante p, soit l'événement contraire F, *est celui où la fréquence relative de* E *est égale à sa probabilité mathématique* p *ou s'en rapproche le plus possible.* En langage plus concis, on peut dire : *La fréquence relative la plus probable d'un événement aléatoire soumis à une longue série d'épreuves identiques est celle qui se rapproche le plus de sa probabilité mathématique.*

72. — *Problème 5.* — On fait une série de 400 tirages d'une urne $U(30, 20 ; 50)$ avec les conventions du numéro 33, définition 8. Quel en est le résultat le plus probable ? *Solution.* On a ici $p = 30 : 50 = 3/5$, $q = 20 : 50 = 2/5$, $r = 400$. Le théorème du numéro 71 donne immédiatement

$$\begin{aligned} k = p.r = {}^3/_5.400 &= 240 \\ n = q.r = {}^2/_5.400 &= 160. \\ \text{Preuve : } k + n = r = &\quad 400 \end{aligned}$$

On est ainsi dispensé de développer $(3/5 + 2/5)^{400}$, comme l'exigerait la méthode générale du numéro 65.

Réponse. — Le résultat le plus vraisemblable est l'apparition de 240 fois boule blanche et 160 fois boule noire, dans un ordre de succession quelconque non prescrit d'avance.

Problème 6. — De la même urne $U(30, 20 ; 50)$, on fait une série de 623 tirages dans les mêmes conditions. Quel en est le résultat le plus probable ? *Solution.* On a maintenant

$$p.r = \frac{3}{5}.623 = 373{,}8$$

qui n'est pas un nombre entier. Il faut donc prendre le plus grand entier contenu dans $p.r + p = 373{,}8 + 0{,}6$; c'est

$$En(374{,}4) = 374 = k ;$$

$$\text{puis } En(q.r + q) = En(249{,}2 + 0{,}4) = 249 = n ;$$

d'où cette *Réponse* : le résultat le plus probable est

$k = 374$ fois boule blanche
$n = 249$ » » noire
$r = 623$

} dans un ordre de succession quelconque non prescrit d'avance.

Problème 7. — De la même urne $U(30, 20, 50)$ et dans les mêmes conditions que ci-dessus, on fait une série de 624 tirages. Quel est maintenant le résultat le plus probable ? *Solution.* On a $p.r = 624.3/5 = 374{,}4$ qui n'est pas un nombre entier. Mais

$$p.r + p = 374{,}4 + 0{,}6 = 375$$

est entier, ainsi que $p.r - q = 374$. Le nombre de fréquence k, et par conséquent aussi $n = 624 - k$, est déterminé bivoquement (v. n° 70, 2°) ; il y a donc deux résultats qui ont la même probabilité mathématique de se réaliser, à savoir :

ou bien	ou bien
$k = 375$ fois boule blanche,	$k = 374$ fois boule blanche,
$n = 249$ » » noire,	$n = 250$ » » noire,
$r = 624$.	$r = 624$.

Problème 8. — D'une urne $U(1, 9 ; 10)$, on fait α) une série de $r_1 = 8$ tirages ; β) une série de $r_2 = 9$ tirages ; γ) une série de $r_3 = 800$ tirages, toujours dans les conditions exposées au numéro 33, définition 8. Quels sont les résultats les plus probables de ces tirages au sort ?

Solution ad α). On a ici $p = 1/10$, $q = 9/10$, $r = 8$, de sorte que $p.r = 0{,}8$ n'est pas un nombre entier ; il faut dès lors (v. n° 71, III 26) prendre le plus grand entier contenu dans $p.r + p = 0{,}9$; c'est zéro ; d'ailleurs le seul nombre entier compris entre $p.r + p = 0{,}9$ et $p.r - q = -0{,}1$ est effectivement zéro ; il s'ensuit $k = 0$ et $n = r = 8$. Le terme maximal du développement de $(0{,}1 + 0{,}9)^8$ est le dernier (v. n° 68, *e*). *Réponse ad* α) : Le résultat le plus probable de cette série est l'apparition de boule noire 8 fois de suite ; sa probabilité mathématique est $= (9/10)^8 = 0{,}43046721$.

Solution ad β). Avec $r = 9$, on a $p.r = 0{,}9$ d'où $p.r + p = 1$ et $p.r - q = 0$. On est dans le cas du n° 70, 2°) ; voir le problème 7. Le développement de $(0{,}1 + 0{,}9)^9$ présente effectivement deux termes maximaux égaux entre eux ; ce sont les deux derniers. *Réponse ad* β) : Le résultat le plus probable est aussi bien l'apparition de boule noire 9 fois de suite, que de boule noire 8 fois et

de boule blanche 1 fois, dans un ordre de succession quelconque. Les deux combinaisons ont la même probabilité mathématique, savoir 0,387 420 489, puisque (v. n° 68, *e*, théorème 2)

$$\left(\frac{9}{10}\right)^9 = 9 \cdot \left(\frac{9}{10}\right)^8 \cdot \left(\frac{1}{10}\right)^1 = \frac{9^9}{10^9} = 0{,}387\,420\,489.$$

Solution ad γ). Avec $r = 800$, on a $p.r = 0{,}1 \cdot 800 = 80 =$ nombre entier (v. n° 70, 1°) et l'on peut de suite donner la *Réponse ad* γ): Le résultat le plus probable est 80 fois boule blanche et 720 fois boule noire, dans un ordre de succession quelconque non fixé d'avance; pour la probabilité mathématique de cette combinaison, voir le numéro 75, problème 11.

La probabilité mathématique du résultat le plus probable

73. — Envisageons de nouveau une longue série de r épreuves identiques dont chacune amène soit l'événement aléatoire E de probabilité mathématique constante $p(E) \equiv p$, soit l'événement contraire F de probabilité mathématique $p(F) \equiv q = 1 - p$. On a vu que T_0, terme maximal dans le développement de $(p+q)^r$, donne la probabilité mathématique du résultat le plus probable de cette série (v. n° 65). Substituons dans l'expression de T_0 donnée par (III 13) les valeurs $k = p.r$ et $n = q.r$, c'est-à-dire les nombres de fréquence les plus probables; voir numéro 70, formule (III 23). La dite expression devient

$$T_0 = \frac{r!}{(p.r)!.(q.r)!} \cdot p^{pr} \cdot q^{qr} \qquad \text{(III 27)}$$

La formule de Stirling permet de poser, avec une approximation suffisante,

$$r! = r^r \cdot e^{-r} \cdot \sqrt{2\pi r}$$
$$(p.r)! = (p.r)^{pr} \cdot e^{-pr} \cdot \sqrt{2\pi pr}$$
$$(q.r)! = (q.r)^{qr} \cdot e^{-qr} \cdot \sqrt{2\pi qr}.$$

Substituées dans (III 27), ces expressions donnent, après une réduction facile,

$$T_0 = \frac{1}{\sqrt{2\pi pqr}}. \qquad \text{(III 28)}$$

Telle est, approximativement, la probabilité mathématique du résultat le plus probable de la série en question.

Corollaires. — On démontre 1) que la valeur de T_0 donnée par (III 28) est toujours approchée par excès ; 2) que l'approximation est en général bonne quand le nombre d'épreuves r est grand ; 3) que l'approximation est très bonne, lorsque le produit pqr est un grand nombre ([52]). *L'erreur relative* est alors inférieure à $\frac{1}{11\,pqr}$.

74. — De la formule (III 28), on déduit immédiatement les conséquences suivantes.

I) *La probabilité mathématique* T_0 *du résultat le plus probable d'une longue série d'épreuves identiques est inversement proportionnelle à la racine carrée du nombre d'épreuves.* Quand on exécute des séries de plus en plus étendues, mettons

$$4,\ 9,\ 16,\ 25,\ \ldots\ldots,\ n^2 \text{ fois plus longues,}$$

la probabilité mathématique T_0 du résultat le plus probable décroît ; elle devient

$$2,\ 3,\ 4,\ 5,\ \ldots\ldots,\ n \text{ fois plus petite.}$$

Voir deux exemples au numéro 75, problèmes 11 et 12.

II) Le nombre d'épreuves r étant donné et constant, la probabilité mathématique du résultat le plus probable devient minimale pour $p = q = \frac{1}{2}$. On peut donc augmenter la dite probabilité T_0 en donnant à l'événement étudié une probabilité mathématique de plus en plus petite, ou au contraire de plus en plus voisine de 1 ; voir pour exemples le numéro 75, problèmes 9 et 10. Il convient cependant de ne jamais oublier que la formule (III 28) n'est qu'approximative et peut donner des valeurs très erronées dans des cas exceptionnels, par exemple quand p est si petit que le produit $2pqr$ est égal ou même inférieur à 1 (v. n° 30).

III) L'unité d'écart absolu étant $u = \sqrt{2pqr}$ (v. n° 57, défini-

([52]) Voir Guido Castelnuovo, *Calcolo delle probabilità* (46), p. 84.

tion 6), on voit que le terme maximal peut se mettre sous la forme

$$T_0 = \frac{1}{\sqrt{2\pi pqr}} = \frac{1}{\sqrt{\pi}} \cdot \frac{1}{\sqrt{2pqr}} = \frac{1}{u\sqrt{\pi}} = \frac{h}{\sqrt{\pi}} \qquad \text{(III 29)}$$

où $h \equiv 1 : u$ représente la mesure de « la précision de la série » (v. n° 77 et n° 87).

Nous allons vérifier les conséquences sus-mentionnées sur quelques exemples.

75. — *Problème 9.* — Dans une série de 1000 épreuves à pile ou face, quelle est la probabilité mathématique du résultat le plus probable ? *Solution* (Voir le n° 38). Comme on a ici $p = q = \frac{1}{2}$, le résultat le plus probable est 500 fois pile et 500 fois face, et sa probabilité mathématique est donnée avec une grande approximation, puisque $pqr = \frac{1}{2} \cdot \frac{1}{2} \cdot 1000 = 250$ n'est pas un petit nombre, par

$$T_0 = \frac{1}{\sqrt{2\pi \cdot 1/2 \cdot 1/2 \cdot 1000}} = \frac{\sqrt{5}}{50\sqrt{\pi}} = 0{,}025\,231.$$

Problème 10. — D'une urne $U(9, 1 ; 10)$, on fait une série de 1000 tirages, dans les conditions du numéro 33, définition 8. Quelle est la probabilité mathématique du résultat le plus probable ? *Solution.* On a ici $p = 9/10$, $q = 1/10$ et $r = 1000$, d'où $k = p.r = 900$ et $n = q.r = 100$. Le résultat le plus probable est donc 900 fois boule blanche + 100 fois boule noire, dans un ordre de succession quelconque non fixé d'avance ; et la probabilité mathématique de ce résultat est à très peu près donnée par

$$T_0 = \frac{1}{\sqrt{2\pi \cdot 0{,}9 \cdot 0{,}1 \cdot 1000}} = \frac{\sqrt{5}}{30\sqrt{\pi}} = 0{,}042\,052.$$

Elle est 5/3 fois plus grande que dans l'exemple précédent.

Remarque. — Comme le nombre d'épreuves est le même dans ces deux exemples, $r = 1000$, on a là une vérification du théorème disant que la probabilité mathématique T_0 est minimale quand

$$p = q = \frac{1}{2}.$$

Problème 11. — Reprenons l'urne $U(1, 9; 10)$ du problème 8 du numéro 72. α) Le résultat le plus probable d'une série de 8 tirages dans les dites conditions est 8 fois de suite boule noire. La formule approximative donne, pour sa probabilité mathématique,

$$T_0 = \frac{1}{\sqrt{2\pi . 0{,}1 . 0{,}9 . 8}} = \frac{5}{6\sqrt{\pi}} = \quad 0{,}470\ 158$$

La valeur exacte en est 0,430 467

Erreur absolue . . 0,039 691 par excès.

Cela correspond à une *erreur relative* de

$$3\ 969\ 100 : 430\ 467 = 9{,}2\ \%.$$

L'erreur est si grande, parce que le produit $pqr = 0{,}1 . 0{,}9 . 8 = 0{,}72$ est un petit nombre.

γ) Le résultat le plus probable d'une série de 800 tirages est 720 fois boule noire + 80 fois boule blanche, dans un ordre de succession quelconque non fixé d'avance ; et sa probabilité mathématique est donnée approximativement par

$$T_0 = \frac{1}{\sqrt{2\pi . 0{,}1 . 0{,}9 . 800}} = \frac{1}{12\sqrt{\pi}} = 0{,}047\ 0158.$$

Comme $pqr = 0{,}1 . 0{,}9 . 800 = 72$, *l'erreur relative* est inférieure à 0,0012, soit à 0,12 % ou 1,2 ‰. A remarquer encore la vérification du théorème énoncé au numéro 74, I.

Problème 12. — D'une urne U (20, 10 ; 30), on fait une série α) de 12 tirages ; β) de 1200 tirages, dans les conditions du n° 33, définition 8. Déterminer les probabilités mathématiques des résultats les plus probables et vérifier le théorème du numéro 74, I.

Solution ad α). Ce problème, déjà traité au numéro 66 (problème 2), va permettre d'éprouver l'approximation de la formule (III 28). Elle donne, pour la probabilité mathématique du résultat le plus probable qui est ici 8 fois boule blanche + 4 fois boule noire, la valeur approximative

$$T_0 = \frac{1}{\sqrt{2\pi . 2/3 . 1/3 . 12}} = \frac{\sqrt{3}}{4\sqrt{\pi}} = \quad 0{,}244\ 301$$

Le développement direct donne 0,238 446

Erreur absolue. . . 0,005 855 par excès.

L'erreur relative est donc 0,005855 : 0,238446 = 0,0246 ou 2,46 %. C'est que le produit $pqr = \frac{2}{3} \cdot \frac{1}{3} \cdot 12 = \frac{8}{3} = 2{,}6\ldots$ est loin d'être un grand nombre.

Solution ad β). Comme $p.r = \frac{2}{3} \cdot 1200 = 800$, le résultat le plus probable est 800 fois boule blanche + 400 fois boule noire, et sa probabilité mathématique est donnée approximativement par

$$T_0 = \frac{1}{\sqrt{2\pi \cdot 2/3 \cdot 1/3 \cdot 1200}} = \frac{\sqrt{3}}{40\sqrt{\pi}} = \quad 0{,}0\ 24\ 430$$

Le calcul exact donne (53) 0,0 24 424

Erreur absolue 0,0 00 006

L'erreur relative, 1/4 ‰ environ, est négligeable ; c'est que le produit $pqr = \frac{2}{3} \cdot \frac{1}{3} \cdot 1200 = 266{,}6\ldots$ n'est pas un petit nombre.

(53) Voir E. Czuber, *Wahrscheinlichkeitsrechnung* (46), p. 135.

CHAPITRE IV

LE THÉORÈME DE JACQUES BERNOULLI ET LA THÉORIE DES ÉCARTS

I. — *Probabilité mathématique que la fluctuation d'une série d'épreuves semblables prenne une valeur donnée d'avance.*

76. — Considérons une longue série de r épreuves identiques dont chacune amène soit l'événement aléatoire E de probabilité mathématique constante $p(E) \equiv p$, soit l'événement contraire F de probabilité mathématique $p(F) \equiv q = 1 - p$. Le premier but de ce chapitre est de déterminer la probabilité mathématique P_{-l}^{+l} que la fluctuation de la série ne dépasse pas des limites fixées d'avance, mettons $\pm l$. A cet effet, recourons au théorème ci-dessus démontré (voir le numéro 63) et appliquons la méthode exposée au numéro 65, question B. Dans le développement de $(p+q)^r$, le terme maximal, T_0, donne la probabilité mathématique de la combinaison la plus probable de notre série, savoir k fois l'événement E et $(r-k)$ fois l'événement F, où $k = pr$ ou $= En(pr+p)$ (v. n° 56, définition 5). Désignons les termes qui se trouvent à gauche de T_0 par T_1, T_2, ..., T_l, ... et ceux à droite de T_0 par T_{-1}, T_{-2}, ..., T_{-l}, ... Le développement aura, dans le voisinage du terme maximal, la forme suivante :

$$(p+q)^r = \ldots + T_l + T_{l-1} + \ldots + T_2 + T_1 + T_0 + T_{-1} + T_{-2} + \ldots + T_{-l+1} + T_{-l} + \ldots \qquad \text{(IV 1)}$$

D'après cela, T_{-l} est la probabilité mathématique que l'écart

absolu de la série soit exactement l, dans l'un seulement des deux sens possibles ; en d'autres termes : T_{-l} est la probabilité mathématique que E se réalise exactement $(k - l)$ fois au cours des r épreuves. De même, T_l est la probabilité mathématique que l'événement étudié arrive exactement $(k + l)$ fois au cours des r épreuves, ou que sa fréquence relative soit $(k + l) : r$. En se basant sur le théorème 1 du numéro 63 et en utilisant la formule de Stirling, on démontre ce

Théorème. — La probabilité mathématique que l'écart absolu d'une série soit exactement x, *ou que son écart relatif soit exactement* x : r, *est donnée approximativement par l'expression*

$$T_x = T_o : e^{\frac{x^2}{2pqr}}. \qquad (IV\ 2)$$

Pour un écart nul, $x = 0$, on retrouve T_o, probabilité mathématique du résultat le plus probable. En introduisant l'unité d'écart absolu, $u = \sqrt{2pqr}$ (v. n° 57), et en tenant compte de l'expression trouvée pour T_o, formule (III 29), on obtient

$$T_x = T_o : e^{\frac{x^2}{u^2}} = \frac{1}{u\sqrt{\pi}} . e^{-\frac{x^2}{u^2}}. \qquad (IV\ 3)$$

Or, $x : u$ est le nombre qui mesure l'écart absolu x ; c'est donc l'*écart réduit* λ qui mesure aussi l'écart relatif (v. n° 58). Dès lors, la probabilité mathématique T_x que l'écart absolu de la série envisagée soit exactement x, ou que son écart réduit soit $\lambda = x : u$, peut se mettre sous la forme

$$T_x = T_o . e^{-\lambda^2} = \frac{\lambda}{x.\sqrt{\pi}} \ e^{-\lambda^2}. \qquad (IV\ 4)$$

Remarque. — Dans la démonstration de la formule (IV 2), on fait intervenir les hypothèses suivantes : 1) que le nombre d'épreuves, r, est grand, au moins 10 000 par exemple ; 2) que l'écart absolu, x, est au plus de l'ordre de grandeur de $\sqrt{r}$; 3) que p ne diffère pas sensiblement de 0,5. Il se trouve que dans la pratique, on peut s'affranchir de ces hypothèses (v. n° 78).

Corollaire. — La probabilité mathématique que l'écart absolu soit égal à x *en plus ou en moins*, et par conséquent l'écart relatif $\pm x : r$, est donnée par $T_x + T_{-x}$.

77. — *Définition 1.* — Introduisons dès maintenant la grandeur

$$h \equiv \frac{1}{\sqrt{2pqr}} = \frac{1}{u}. \qquad \text{(IV 5)}$$

On l'appelle *la mesure de la précision absolue de la série* ou plus brièvement *la précision de la série*. Ce nom fut introduit dans la théorie des erreurs d'observation ; là, son usage est approprié, car une erreur d'observation n'est autre chose qu'un « écart absolu », savoir la différence entre le résultat d'une mesure effectivement exécutée et la vraie valeur, connue ou non, de la grandeur qu'il s'agit de mesurer : longueur, angle, intervalle de temps etc. Plus tard, on étendit le nom et la notion de « précision d'une série » au problème des épreuves répétées, voyant qu'il conduit aux mêmes formules que celui des erreurs d'observation, quoique ces deux problèmes soient bien différents au point de vue des concepts mis en œuvre. En comparant (IV 5) avec les définitions posées aux numéros 57 et 58, on voit que l'écart réduit devient

$$\lambda \equiv l : u = l.h, \qquad \text{(IV 6)}$$

en paroles : *l'écart réduit est égal au produit de l'écart absolu par la mesure de la précision (absolue) de la série*. L'introduction de h dans la formule (IV 4), où l'écart absolu l est représenté par x, donne

$$T_x = T_o : e^{h^2x^2} = \frac{h}{\sqrt{\pi}} . e^{-h^2x^2}. \qquad \text{(IV 7)}$$

Définition 2. — Comme on distingue entre fréquence absolue et fréquence relative, entre écart absolu et écart relatif etc. (v. les numéros 52, 55, 56, 57), on peut distinguer entre « mesure de la précision absolue » et « mesure de la précision relative », h_r. De même qu'on a dénommé « mesure de la précision absolue » la valeur réciproque de l'unité d'écart absolu, j'entendrai par *mesure de la précision*

relative, et désignerai par h_r, la valeur réciproque de l'unité d'écart relatif ou de l'écart étalon (voir n° 57), de sorte que

$$h_r \equiv \frac{r}{u} = \frac{r}{\sqrt{2pqr}} = \sqrt{\frac{r}{2pq}} \cdot \qquad \text{(IV 8)}$$

Nous aurons lieu d'y revenir (voir les numéros 98 et 99).

78. — La propriété la plus utile de la fonction T_x définie par (IV 7) est de décroître très rapidement quand x augmente. Cette décroissance est même si rapide que l'on peut abandonner l'hypothèse qui sert de base à la démonstration de la formule, savoir que x soit au plus de l'ordre de grandeur de $\sqrt{r}$. Et quand $\sqrt{pqr}$ est un grand nombre, on peut étendre l'expression approchée de la probabilité mathématique donnée par T_x à n'importe quelle valeur de x. L'erreur commise reste négligeable.

Exemples. — D'une urne U (15, 10 ; 25), on fait une série de 2500 tirages dans les conditions exposées au numéro 33, définition 8. Puisqu'ici $p = 15/25 = 3/5$, $q = 10/25 = 2/5$, $r = 2500$, on a $pr = 1500$, $qr = 1000$. Dès lors :

1) Le résultat le plus probable est 1500 fois boule blanche + 1000 fois boule noire.

2) La probabilité mathématique en est très approximativement

$$T_0 = \frac{1}{\sqrt{2\pi . 0{,}6\ 0{,}4 . 2500}} = \frac{\sqrt{3}}{60\sqrt{\pi}} = 0{,}016\,2868.$$

3) La probabilité mathématique d'un écart absolu de — 50, c'est-à-dire la probabilité mathématique que boule blanche sorte juste 1500 — 50 = 1450 fois et par conséquent boule noire juste 1050 fois, dans un ordre de succession quelconque non prescrit d'avance, est donnée par (IV 2), quand on y substitue $x = -50$; on trouve pour l'exposant

$$\frac{x^2}{2pqr} = \frac{2500}{2 . 0{,}6 . 0{,}4 . 2500} = \frac{25}{12}, \text{ de sorte que}$$

$$T_{-50} = T_0 . e^{-\frac{25}{12}} = 0{,}016\,2868 . 0{,}1245145 = 0{,}002\,0279.$$

4) La probabilité mathématique que l'écart absolu x soit exactement 80 ou 100 ou 150 est donnée très approximativement par

$$T_{80} = T_o \cdot e^{-\frac{16}{3}} = 0{,}016\,287 \cdot 0{,}004\,828\,0 = 0{,}000\,078\,633.$$

$$T_{100} = T_o \cdot e^{-\frac{25}{3}} = 0{,}016\,287 \cdot 0{,}000\,240\,6 = 0{,}000\,003\,918\,6.$$

$$T_{150} = T_o \cdot e^{-\frac{75}{4}} = 0{,}016287 \cdot 0{,}000000\,0071941 = 0{,}000000\,00011717.$$

79. — La fonction T_x définie par (IV 7) est une fonction paire, c'est-à-dire que $T_x = T_{-x}$. Il s'ensuivrait que des écarts égaux mais de sens opposé ont la même chance de se produire. Il convient de rappeler à ce propos que cette propriété n'a lieu rigoureusement que si $p = q = \frac{1}{2}$. Quand $p \neq q$, il y a asymétrie dans le développement de $(p + q)^r$; et $p > q$ entraîne $T_l > T_{-l}$ (v. n° 76, n° 68, *d*, *e*). Mais cette asymétrie reste insensible et l'on peut poser $T_x = T_{-x}$ dans les conditions suivantes : *a*) Si $\sqrt{pqr}$ est un grand nombre ; ou bien *b*) Si l'écart absolu est inférieur en valeur absolue à $2\sqrt{pqr}$ et en même temps la différence $(p - q)$ négligeable par rapport à p. Vérifions ces propriétés sur un exemple.

Exemple. — D'une urne U (6, 24 ; 30), on fait une série de $r = 100$ tirages dans les conditions exposées au numéro 33, définition 8. On a donc ici

$$p = 6/30 = \frac{1}{5}, \quad q = 24/30 = \frac{4}{5}, \quad r = 100,$$

d'où $pr = 20$ et $qr = 80$; dès lors :

1) Le résultat le plus probable est 20 fois boule blanche + 80 fois boule noire, dans un ordre de succession quelconque.

2) La probabilité mathématique de ce résultat est approximativement égale à

$$T_o = \frac{1}{\sqrt{2\pi \cdot 0{,}2 \cdot 0{,}8 \cdot 100}} = \frac{\sqrt{2}}{8\sqrt{\pi}} = \ldots\ldots \quad 0{,}099\,736$$

La valeur exacte en est $\binom{100}{20}\left(\frac{1}{5}\right)^{20}\left(\frac{4}{5}\right)^{80} =$ 0,099 301

Erreur absolue . . . 0,000 435

L'erreur relative $= 435 : 99301 = 0{,}004\,380\ldots$, soit $4\,{}^0/_{00}$ du résultat exact, est négligeable.

3) Quelle est la probabilité mathématique d'un écart absolu de 10, c'est-à-dire 30 fois boule blanche + 70 fois boule noire, et celle d'un écart absolu de (— 10), c'est-à-dire 10 fois boule blanche + 90 fois boule noire ? La formule (IV 2) donne pour les deux la même valeur, savoir

$$T_{10} = T_{-10} = \frac{\sqrt{2}}{8\sqrt{\pi}} . e^{-\frac{25}{8}} = 0{,}099\,736 . 0{,}043\,9369 = 0{,}004\,3821,$$

tandis que le calcul exact donne

T_{10}	$=$	0,005 19	et T_{-10}	$=$ 0,003 36
		0,004 38		0,004 38
Erreurs absolues :		0,000 81		— 0,001 02

Ici donc, la formule (IV 2) donne une approximation insuffisante, puisque l'erreur relative va jusqu'à $102 : 336 = 0{,}303\ldots$, soit plus de 30 ${}^0/_0$ du résultat exact. Cela provient de ce que les conditions sus-mentionnées ne sont pas remplies, car 1° $\sqrt{pqr} = \sqrt{0{,}2 . 0{,}8 . 100} = 4$ n'est pas un grand nombre ; 2° l'écart absolu, 10, surpasse $2\sqrt{pqr} = 8$; et 3° la différence $(p - q) = - 0{,}6$ n'est pas négligeable en comparaison de $p = 0{,}2$.

Remarque. — M^r^ Castelnuovo ([46]) fait observer (p. 72) que la valeur $T_x = 0{,}00438$ donnée par la formule approximative (IV 2) est très proche de la moyenne arithmétique 0,00428 des valeurs exactes de T_{10} et de T_{-10}. On démontre que ce fait est général.

II. — *Probabilité mathématique que l'écart d'une série d'épreuves répétées reste compris entre des limites fixées d'avance.*

80. — Dans les applications, on n'a presque jamais à s'occuper de la probabilité mathématique que l'écart ou la fluctuation d'une série soit exactement $= l$; cette probabilité devient toujours très

petite quand le nombre d'épreuves, r, devient très grand (v. n° 74, I et n° 35). Par contre, il est important de savoir avec quelle probabilité on peut s'attendre à ce que l'écart reste inférieur à certaines limites fixées d'avance, par exemple à ce que la fluctuation ne dépasse pas l, ni en plus ni en moins. C'est la question B soulevée au numéro 65. Or, la probabilité mathématique que l'écart absolu soit

nul ... est $= T_0$,
exactement 1, en plus ou en moins, est $= T_1 + T_{-1}$,
» 2, » » , » $= T_2 + T_{-2}$,
» 3, » » , » $= T_3 + T_{-3}$,
. . . . . .
» l, » » , » $= T_l + T_{-l}$.

Dès lors, la probabilité mathématique P_{-l}^{+l} que l'écart absolu ne dépasse pas $\pm l$ est la somme de tous ces termes (v. n° 32), de sorte que $P_{-l}^{+l} = T_l + T_{l-1} + \ldots\ldots + T_1 + T_0 + T_{-1} + \ldots\ldots + T_{-l}$, où les T_a sont donnés par les expressions (IV 2) ou (IV 4) du numéro 76. On réussit à déterminer la somme de cette série à l'aide d'une célèbre formule sommatoire très générale due à L. Euler. Il faut introduire une nouvelle fonction transcendante, $\Theta(x)$, fondamentale pour toute la théorie des écarts. On l'appelle *la transcendante de Kramp-Laplace.*

Définition 3. — *La fonction Théta* est donnée ordinairement sous forme d'intégrale définie :

$$\Theta(x) \equiv \frac{2}{\sqrt{\pi}} \cdot \int_0^x e^{-x^2} \cdot dx. \qquad \text{(IV 9)}$$

Vu la grande importance pratique de cette fonction Théta dans plusieurs domaines, on en a calculé, comme pour la fonction logarithmique, des tables étendues dont la plus ancienne est celle dressée par Kramp (54) pour une théorie des réfractions, en 1798. Une de

(54) Chr. Kramp, *Analyse des réfractions astronomiques et terrestres*, Strasbourg et Leipzig, an VII, corrigée et reproduite par A. de Morgan, *Theory of probabilities*, Encyclop. Metrop. II, 1845. Des tables très pré-

ces tables est reproduite à la fin du présent volume. Voici la deuxième partie de l'important théorème de Jacques Bernoulli (v. n° 71) démontré par Laplace dans sa *Théorie analytique des probabilités* (v. n° 11).

Théorème fondamental de Jacques Bernoulli (deuxième partie). — Soit une longue série de r épreuves identiques exécutées dans les mêmes conditions, chacune amenant soit l'événement aléatoire E soit son contraire F, leurs probabilités mathématiques respectives $p(E) \equiv p$ et $p(F) \equiv q = 1 - p$ restant constantes. Dans ces conditions, *la probabilité mathématique* P_{-l}^{+l} *que l'écart absolu de la série ne dépasse pas* $\pm l$ *et que par conséquent l'écart relatif ne dépasse pas* $\pm l : r$, *en d'autres termes, la probabilité mathématique que le nombre de fréquence de l'événement* E *tombe entre les limites* pr $-$ l *et* pr $+$ l, *est donnée très approximativement par*

$$P_{-l}^{+l} = \Theta(\lambda) + \frac{h}{\sqrt{\pi}} \cdot e^{-\lambda^2} \qquad \text{(IV 10)}$$

où $\lambda = h.l = l : \sqrt{2pqr}$ *est l'écart réduit de la série (v. n° 58) et où* Θ *représente la fonction transcendante de Kramp-Laplace* définie par la formule (IV 9) (Voir la table à la fin du présent volume).

Remarque. — Le premier terme de la formule (IV 10) donne une valeur très voisine de 1 dès que l'écart réduit, λ, dépasse 3. Pour $\lambda = 4$, la différence $1 - \Theta(\lambda)$ est de l'ordre du cent-millionième. Dans ces cas, la contribution du second terme du deuxième membre de (IV 10) est tout-à-fait négligeable. Même pour des valeurs plus petites de l'écart réduit, le premier terme, $\Theta(\lambda)$, est

cises de $\theta(x)$ ont été publiées par R. Radau dans les *Annales de l'Observat.* Paris, *Mémoires*, t. 18, 1885, mém. n° 2 et par A.-A. Markoff, *Table des valeurs de l'intégrale* $\int_x^\infty e^{-t^2} dt$, Saint-Pétersbourg, 1888. On en trouve également chez Joseph Bertrand (49), A.-A. Markoff (51), E. Czuber (46) et (53), H. Bruns (40), etc. Voir aussi H. Opitz, *Die Kramp-Laplacesche Transcendente und ihre Umkehrung*, Osterprogramm 1900 des Kœnigstædt. Realgymnasiums zu Berlin.

de beaucoup prépondérant. Dans les applications numériques, on peut presque toujours faire abstraction du second terme et prendre

$$P_{-l}^{+l} = \Theta(\lambda) \qquad \text{(IV 11)}$$

Nous allons signaler quelques corollaires du théorème de Jacques Bernoulli, appelé aussi « loi des grands nombres » par quelques auteurs. Ces corollaires, démontrés rigoureusement par Laplace (v. n° 11), dépassent ce que Jacques Bernoulli lui-même avait établi. Ils mettent en lumière le sens de ce théorème fondamental et font entrevoir sa grande portée.

81. — α) Supposons que la probabilité mathématique P_{-l}^{+l} exprimée par (IV 10) soit donnée et fixe. *Alors les limites* $\pm l$ *de l'écart absolu ou de la fluctuation de la série augmentent proportionnellement à la racine carrée du nombre d'épreuves r.* Donc, si l'on fait des séries d'épreuves de plus en plus longues et de plus en plus nombreuses, il faut s'attendre avec une probabilité mathématique donnée P (que l'on peut choisir aussi près de 1 que l'on voudra) à ce que l'écart absolu finisse par dépasser toute limite fixée d'avance, si grande soit-elle. Au fur et à mesure qu'on allonge les séries d'épreuves et qu'on en multiplie le nombre, il devient de moins en moins probable que l'écart absolu reste inférieur à un nombre N choisi d'avance. En d'autres termes, si l'on augmente indéfiniment le nombre des épreuves, on doit s'attendre, avec une probabilité qui frise la certitude, à ce qu'il finira par se réaliser une série dont la fluctuation dépassera N, quelque grand que l'on ait choisi N à l'avance. Voir n° 89, définition 10.

β) Supposons de nouveau la probabilité mathématique P_{-l}^{+l} donnée et fixe. *Alors les limites* $\pm l : r$ *de l'écart relatif se resserrent de plus en plus, proportionnellement à la racine carrée du nombre d'épreuves r.* Si donc on fait sur le même événement aléatoire des séries d'épreuves de plus en plus longues et de plus en plus nombreuses, il faut s'attendre avec une probabilité mathématique donnée (que l'on peut choisir aussi voisine de 1 que l'on voudra) à ce que l'écart relatif finisse par devenir plus petit, en valeur absolue, que toute limite fixée d'avance. Au fur et à mesure que l'on allonge les séries d'épreuves et qu'on en multiplie le nombre, il

devient de moins en moins probable que l'écart relatif d'une série dépasse un nombre positif ε choisi d'avance arbitrairement petit.

γ) Supposons effectuée, sur un même événement aléatoire E dont la probabilité mathématique p reste constante, une série de r épreuves identiques et que la limite de l'écart relatif, mettons L, soit donnée et fixe. En vertu de (IV 10), la probabilité mathématique P_{-L}^{+L} que l'écart relatif de la série ne dépasse pas L dépend du nombre d'épreuves r, puisque l'écart réduit, λ, en dépend (v. n° 58). Or, *le théorème de Jacques Bernoulli dit que*, L *restant fixe*, $P_{-L}^{+L} \to 1$ *quand* $r \to \infty$; c'est-à-dire que l'on peut prendre le nombre d'épreuves r suffisamment grand pour que la quantité positive ε étant choisie d'avance arbitrairement petite, la différence $1 - P_{-L}^{+L}$ finisse par devenir encore plus petite que ε ; en d'autres termes : il est possible de prendre r tellement grand qu'on ait la quasi-certitude (c'est-à-dire une probabilité mathématique aussi voisine que l'on voudra de 1) que la fréquence relative de l'événement étudié différera aussi peu que l'on voudra de sa probabilité mathématique p connue ou non. *Si donc on augmente indéfiniment le nombre des épreuves, on peut être pratiquement certain qu'il finira par se réaliser une série dont l'écart relatif sera inférieur à* L *en valeur absolue*, et cela quelque petit que l'on ait choisi d'avance $L > 0$.

δ) Le théorème de Jacques Bernoulli enseigne donc un fait très curieux : si l'on allonge indéfiniment les séries d'épreuves, on doit s'attendre avec une quasi-certitude à ce que l'écart absolu croisse au delà de toute limite, tandis que l'écart relatif tende vers zéro ; autrement dit : *Dans une suite illimitée de séries, de plus en plus longues, d'épreuves répétées dans les mêmes conditions, on a une probabilité mathématique aussi voisine que l'on voudra de 1 pour que l'écart absolu devienne infiniment grand et en même temps l'écart relatif infiniment petit* (v. n° 89).

82. — Répétons encore expressément que le théorème de Jacques Bernoulli ne contient que des propositions relatives à des probabilités mathématiques ; il enseigne à déduire de probabilités mathématiques données, p, q, d'autres probabilités mathém

tiques. En conséquence, il ne doit servir qu'à former des conjectures, à justifier des expectatives, non à prédire comment se dérouleront effectivement les phénomènes physiques. Pour ces prévisions de faits matériels, on est obligé de recourir à un postulat qui est la loi empirique du hasard (v. n° 54). Qu'on effectue matériellement une série d'épreuves et qu'on trouve un écart dépassant très notablement les limites prévues, par exemple qu'on joue à pile ou face une série de 10 000 coups où face se montre plus de 6 000 fois ; on n'aura pas infirmé le théorème, puisque celui-ci est une proposition abstraite qui résulte d'une énumération rigoureuse dont on la déduit par un raisonnement mathématique impeccable ; mais on aura rencontré par hasard un cas extraordinairement rare ; on pourra même en conclure qu'une cause perturbatrice a troublé les effets du hasard, par exemple que la pièce utilisée n'était pas bonne dans le sens du n° 38 (v. la *remarque* du n° 60).

Règle. — 1) Dans les problèmes relatifs aux probabilités mathématiques des écarts ou des fluctuations, on doit commencer par déterminer l'*écart réduit* λ (v. n° 58). Celui-ci connu, on utilise la table des valeurs de la fonction $\Theta(\lambda)$. 2) Le plus souvent, on peut faire abstraction du second terme du deuxième membre de la formule (IV 10) et prendre $P_{-l}^{+l} = \Theta(\lambda)$. 3) Pour une approximation plus grande, on a recours au théorème d'Eggenberger (v. n° 85).

83. — *Problème 1.* — On lance 600 fois un dé à jouer. Quelle est la probabilité mathématique que le 3 par exemple soit amené : α) de 85 à 115 fois ? β) de 70 à 130 fois ? γ) que l'écart relatif de cette série de 600 coups dépasse $\pm$ 8 % ?

Solutions. — La probabilité mathématique de l'apparition du 3 est $p = \frac{1}{6}$ (v. n° 23, problème 1), donc $q = 1 - p = 5/6$. Comme $p.r = \frac{1}{6} . 600 = 100$ et $q . r = 500$, le résultat le plus probable est : 100 fois apparition du point 3.

Question α). — L'écart absolu doit être au plus 115 − 100 = 15. De $l = 15$, on déduit l'écart réduit

$$\lambda = 15 : \sqrt{2 \cdot 1/6 \cdot 5/6 \cdot 600} = 0{,}3 \cdot \sqrt{15} = 1{,}161\,89\ldots$$

On trouve dans la table $\theta(\lambda) = \theta(1{,}1619) = 0{,}89966$, soit presque 0,90. On peut donc parier 9 contre 1 que, dans une série de 600 coups de dé, le point 3 sera amené de 85 à 115 fois.

Question β). — L'écart absolu maximum où la fluctuation est ici 130 − 100 = 30. C'est le double du précédent ; l'écart réduit est donc aussi le double, d'où $\lambda = 0{,}6 \cdot \sqrt{15} = 2{,}323\,79$. La table donne $\theta(\lambda) = \theta(2{,}3238) = 0{,}9990$. On peut donc parier 999 contre 1 que l'écart absolu ne dépassera pas ± 30, c'est-à-dire que dans la série en question, le point 3 sera amené de 70 à 130 fois.

Question γ). — A un écart relatif de 8 % = 0,08 correspond ici un écart absolu $l = 600 \cdot 0{,}08 = 48$, d'où un écart réduit de $\lambda = 48 : \sqrt{2 \cdot 1/6 \cdot 5/6 \cdot 600} = 0{,}96 \cdot \sqrt{15} = 3{,}718\,06\ldots$ La probabilité mathématique que l'écart reste compris entre les limites en question est donc $\theta(\lambda) = \theta(3{,}718) = 0{,}999\,999\,853$, et par conséquent celle qu'il dépasse ± 8 % est $1 - \theta(3{,}718) = 0{,}000\,000\,147$, soit de l'ordre du dix-millionième, c'est-à-dire tout à fait négligeable. Pratiquement, on peut donc être certain que l'écart (relatif) d'une série de 600 coups de dé ne dépassera pas ± 8 %.

84. — La question de l'exactitude obtenue par l'emploi de la formule de Laplace est tranchée par le théorème suivant ([55]).

Théorème. — En calculant par la formule de Laplace (IV 10) *la probabilité mathématique* P_{-l}^{+l} *que l'écart absolu d'une série de* r *épreuves ne dépasse pas* ± l, *on commet une erreur relative inférieure à* $\frac{1}{pqr}$ *si l'écart absolu* $l \leqslant 2\sqrt{pqr}$. L'erreur relative reste plus petite que $\frac{5}{pqr}$, si l est compris entre le double et le

(55) Voir G. Castelnuovo, *Calcolo delle probabilità* (46), p. 89.

triple de l'écart quadratique moyen $\mu \equiv \sqrt{pqr}$ (voir les numéros 87 et 95), en d'autres termes, si

$$2\sqrt{pqr} \leqslant l \leqslant 3\sqrt{pqr}.$$

Pour illustrer ce théorème par un exemple, examinons l'expérience exécutée par Buffon qui fit jeter en l'air une pièce de monnaie 4040 fois. Il obtint 2048 fois face, au lieu de $\frac{1}{2} \cdot 4040 = 2020$ qui est le nombre de fréquence le plus probable.

Problème 2. — Dans une série de 4040 coups à pile ou face, quelle est la probabilité mathématique que l'écart absolu ne dépassera pas 28 ? en d'autres termes : que face apparaîtra de 1992 à 2048 fois ?

Solution. — Comme ici $p = q = 1/2$, $r = 4040$ et $l = 28$, on en déduit tout d'abord l'écart réduit

$$\lambda = 28 : \sqrt{2 . 1/2 . 1/2 . 4040} = 14 : \sqrt{505} = 0{,}622\,99.$$

La table donne ensuite en première approximation

$$P_{-28}^{+28} = \Theta(0{,}623) = 0{,}621\,71.$$

Pour tenir compte du second terme de la formule de Laplace, il faudrait ajouter (v. n° 80)

$$\frac{h}{\sqrt{\pi}} \cdot e^{-\lambda^2} = \frac{e^{-(0{,}623)^2}}{\sqrt{2\pi . 1/2 . 1/2 . 4040}} = \frac{0{,}67833}{2\sqrt{505\,\pi}} = 0{,}008\,515$$

$$0{,}621\,710$$

Il s'ensuit que la probabilité cherchée est = $\overline{0{,}630\,225}$

On peut donc parier 63 contre 37 que l'écart absolu ne dépassera pas 28, ou que l'écart relatif ne dépassera pas $28 : 4040 = 0{,}00693$, c'est à-dire qu'il restera inférieur à 0,007 ou 7 pour mille. Si l'on admet que seul le hasard soit la cause de l'écart constaté, il n'y a donc pas lieu de s'étonner. Néanmoins, Poisson a cherché la probabilité mathématique que la pièce utilisée par Buffon fût irrégulière, donnant à l'arrivée de face une probabilité mathématique plus grande qu'à celle de pile. C'est un problème qui rentre dans

le chapitre dit de la probabilité des causes ; Poisson l'a résolu par le théorème de Bayes et a trouvé 0,810 43.

Remarque I. — En faisant abstraction du second terme de la formule de Laplace et utilisant (IV 11) au lieu de (IV 10), on commet donc, dans ce cas particulier, une erreur relative de 8515 sur 630 225 ou de 1,35 % ou un et tiers pour cent.

Remarque II. — L'écart constaté par Buffon n'a rien d'extraordinaire, même pour une pièce de monnaie rigoureusement bonne dans le sens du n° 38 ; on a en effet la probabilité mathématique $1 - P_{-28}^{+28} = 1 - 0{,}630 = 0{,}37$ pour l'événement contraire, c'est-à-dire pour que la fluctuation dépasse 28. Si donc on faisait un très grand nombre de séries d'épreuves, chacune comprenant 4040 jets d'une pièce parfaite dans le sens du n° 38, on devrait s'attendre à trouver, en moyenne, 37 séries sur 100, où le nombre de fréquence de face différerait de 2020 au moins autant que dans le cas étudié par Buffon.

85. — Pour tenir compte du second terme de la formule de Laplace (IV 10), on peut chercher à prendre, au lieu de l'écart réduit λ, une valeur légèrement supérieure λ' telle que

$$\Theta(\lambda') = \Theta(\lambda) + \frac{h}{\sqrt{\pi}} \cdot e^{-h^2 l^2}$$

J. Eggenberger, qui a eu cette idée, a démontré qu'il convient de prendre

$$\lambda' = \lambda + \frac{1}{2\sqrt{2pqr}} = \frac{l + 1/2}{u}, \qquad \text{(IV 12)}$$

ce qui revient à augmenter de 1/2 les limites de l'écart absolu l ou de $\frac{1}{2r}$ les limites de l'écart relatif (56). De cette façon, on a l'avantage de pouvoir, tout en évitant de longs calculs supplémentaires,

(56) J. Eggenberger, *Beiträge zur Darstellung des Bernoullischen Theorems, der Gammafunktion und des Laplace'schen Integrals.* Inauguraldissertation, Bern 1893. Voir aussi Berner Mitteilungen 50 (1894) et Zeitschrift für Mathem. und Physik 45 (1900), p. 43.

augmenter notablement l'exactitude du résultat en utilisant la même table de $\theta(\lambda)$. La formule (IV 12) résume *le théorème d'Eggenberger.*

Exemple. — Appliquons ce théorème au cas étudié par Buffon (v. n° 84, problème 2). On calcule

$$\lambda' = \frac{28 + 0{,}50}{\sqrt{2.1/2.1/2.4040}} = \frac{14{,}25}{\sqrt{505}} = 0{,}634\,12.$$

On trouve dans la table de $\theta(\lambda)$, par interpolation,

$$P_{-28}^{+28} = \theta(0{,}634\,12) = 0{,}630\,187$$

soit 0,000 038 de moins qu'en faisant le calcul rigoureux.

Pratiquement, la formule d'Eggenberger corrige l'erreur provenant du fait de négliger le second terme dans la formule de Laplace (IV 10).

86. — La formule de Laplace permet de résoudre encore une autre catégorie de *Problèmes.* — Combien faut-il faire d'épreuves pour pouvoir s'attendre, avec une probabilité mathématique donnée b, à ce que α) l'écart absolu d'une série de r épreuves identiques ne dépasse pas des limites fixées d'avance, par exemple $\pm l$? ou bien à ce que β) l'écart relatif de la série reste compris dans un intervalle donné, par exemple $< -\varepsilon, +\varepsilon >$?

Solutions. — Pour simplifier, négligeons le second terme de la formule de Laplace. Elle se réduit alors à (IV 11), savoir $P_{-l}^{+l} = \theta(\lambda)$ et l'on peut procéder comme suit.

α) Le premier membre de (IV 11) est donné $= b$, donc connu ; la relation $\theta(\lambda) = b$ permet de trouver le λ correspondant à b, grâce aux tables Théta, comme une table de logarithmes ordinaires permet de calculer le nombre correspondant à un logarithme donné. L'écart réduit λ trouvé, on recourt à la formule de définition $\lambda = l : \sqrt{2pqr}$ dans laquelle λ, l, p, q, sont maintenant connus. On peut donc en déduire la valeur cherchée de r. De $\lambda^2 2pqr = l^2$, on tire en effet

$$r = \frac{l^2}{2pq\lambda^2} \qquad \text{(IV 13)}$$

β) Dans le cas où ce sont les limites $\pm \varepsilon$ de l'écart relatif qui sont données, il faut se rappeler la définition 4 du n° 56; elle donne $l = \varepsilon . r$ et l'équation ci-dessus devient $\lambda^2 2pqr = \varepsilon^2 r^2$. On en déduit, après simplification par r, la formule

$$r = \frac{2\lambda^2 pq}{\varepsilon^2} \qquad \text{(IV 14)}$$

Problème 3. — D'une urne $U(25, 5; 30)$, combien faut-il faire de tirages dans les conditions du n° 33, définition 8, pour pouvoir parier 999 contre 1 que la fréquence relative de boule blanche différera de sa valeur la plus probable, α) de 5 % au plus? β) de 1 % au plus? γ) de 1 ‰ au plus?

Solutions. — Parier 999 contre 1 revient à la probabilité mathématique $P_{-l}^{+l} = 0{,}999$. L'égalité $\theta(\lambda) = 0{,}999$ donne, par l'emploi de la table de la fonction Théta, $\lambda = 2{,}3268$. C'est l'écart réduit de la série (v. la règle du n° 82).

Nous avons maintenant $p = 25 : 30 = 5/6$ et $q = 1/6$. Dès lors, la valeur la plus probable de la fréquence relative est 5/6, dont α) le 5 % est $= \frac{5}{6} . \frac{5}{100} = \frac{1}{24}$; et dont β) le 1 % est $= \frac{5}{6} . \frac{1}{100} = \frac{1}{120}$; et dont γ) le 1 ‰ est $= \frac{1}{1200}$. La formule (IV 14) conduit alors aux résultats suivants.

Question α). — On trouve

$$r_1 = \frac{2 . 2{,}3268^2 . 5 . 1 . 24^2}{6 . 6 . 1} = 160 . 2{,}3268^2 = 866{,}24 \text{ soit } 867 \text{ épreuves.}$$

Question β). — On trouve

$$r_2 = \frac{2 . 2{,}3268^2 . 5 . 1 . 120^2}{6 . 6} = 4\,000 . 2{,}3268^2 = 21\,655{,}993$$

soit 21656 épreuves.

Question γ). — Le 1 pour mille étant 10 fois moindre que le 1 pour cent, on trouve $r_3 = 10^2 . r_2 = 2\,165\,600$ épreuves, ce que confirmerait la formule (IV 14). Si donc on fait 2 165 600 tirages

de l'urne U (25, 5; 30), on pourra parier 999 contre 1 que la fréquence relative de boule blanche différera au plus de 1 pour mille de sa valeur la plus probable qui est ici 5/6.

III. — *L'écart quadratique moyen d'une série d'épreuves répétées.*

87. — Lorsqu'on veut comparer entre elles diverses séries d'observations ou d'épreuves relatives à des événements fortuits, séries pouvant se rapporter à des problèmes différents, il est utile d'introduire certains paramètres qui permettent de se faire une idée de la grandeur des écarts comparée à la longueur de la série. L'un de ces paramètres, h, déjà mentionné aux numéros 74 et 77, est la mesure de la précision de la série (v. les numéros 98 et 99) :

$$h \equiv 1 : \sqrt{2pqr}, \text{ où } p + q = 1. \qquad \text{(IV 5)}$$

Définition 4. — Un autre paramètre très important est *l'écart quadratique moyen de la fréquence absolue*, ou simplement *l'écart quadratique moyen*, μ; c'est la *standard deviation* des statisticiens anglais, *die mittlere Abweichung* en allemand ([57]), *lo scarto quadratico medio* en italien ([58]). Beaucoup d'auteurs le définissent par la formule

$$\mu = \sqrt{pqr}. \qquad \text{(IV 15)}$$

En réalité, cette formule exprime un théorème remarquable que nous allons démontrer. Rappelons à cet effet que l'écart absolu ou la fluctuation d'une série de r épreuves semblables dont chacune amène nécessairement soit l'événement aléatoire E soit son contraire F, est par définition $l \equiv k - p.r$, où k est le nombre de fréquence réellement observé de E (v. n° 56). Ce nombre k pouvant varier de 0 à r, suivant les séries, l'écart absolu l peut manifestement prendre aussi $r + 1$ valeurs différentes, puisque par hypothèse le produit $p.r$ reste constant. L'écart quadratique moyen, μ, se définit alors comme suit.

(57) Voir p. ex. E. Czuber, *l. cit.* (46), p. 144-146.
(58) Voir p. ex. Castelnuovo, *l. cit.* (52), p. 98-99.

Définition 5. — 1° Qu'on forme le carré de chacun des écarts absolus possibles ; 2° qu'on multiplie ce carré par la probabilité mathématique que l'écart a de se réaliser ; 3° qu'on fasse la somme des produits ainsi obtenus ; 4° qu'on prenne enfin la racine carrée de cette somme ; on aura ainsi *l'écart quadratique moyen*, μ. Traduite en formule, cette définition devient :

$$\mu^2 \equiv \sum_{k=0}^{r} (k - pr)^2 . \binom{r}{k} . p^k . (1-p)^{r-k} \qquad \text{(IV 16)}$$

Or, chose remarquable, on a ce curieux

Théorème 1. — *La somme* (IV 16) *se réduit à pqr.* On peut le démontrer de différentes manières dont voici la plus rapide ; elle rappelle un raisonnement de Pascal (v. n° 6). La fluctuation l d'une série de r épreuves est définie par

$$k = p.r + l \qquad (a)$$

Supposons qu'on fasse encore une épreuve ; on aura une série de $(r+1)$ épreuves et le résultat le plus probable en est que l'événement étudié E se réalisera $(r+1).p$ fois (v. n° 70). Or, cette $(r+1)^{\text{ième}}$ épreuve, ou bien elle augmente de 1 le nombre de fréquence k ou bien elle ne le modifie pas du tout, suivant qu'elle amène l'événement E ou son contraire F. Envisageons séparément ces deux possibilités.

1) Si c'est E qui se réalise, E sera arrivé pour la $(k+1)^{\text{ième}}$ fois et l'on aura pour son nombre de fréquence, en vertu de (a) et de $p + q = 1$, l'expression

$$k + 1 = (r+1)p + (l+q).$$

Donc la fluctuation, de l qu'elle était, devient $(l+q)$, c'est-à-dire augmente de q ; et la probabilité mathématique que cela se produise est $p(E) = p$.

2) Si c'est l'événement contraire, F, qui se réalise, le nombre de fréquence k restera inchangé et l'on aura, en vertu de (a),

$$k = (r+1)p + (l-p),$$

ce qui montre que la fluctuation, de l qu'elle était, devient $l - p$, c'est-à-dire diminue de p ; et la probabilité mathématique que cela se produise est $p(F) = q$. Donc, puisqu'on multiplie le carré de chaque fluctuation par la probabilité mathématique qu'elle a de se réaliser, cette $(r + 1)^{\text{ième}}$ épreuve a pour effet d'augmenter la somme en question de

$$p \cdot q^2 + q \cdot (-p)^2 = pq^2 + qp^2 = pq(p + q) = pq.$$

On voit par là que cette augmentation est indépendante de r ; donc chaque épreuve a pour effet d'augmenter de pq la somme en question ; et comme la valeur initiale en est nulle (quand il y a zéro épreuve, $r = 0$, il n'y a pas non plus d'écart), la somme vaut autant de fois pq qu'il y a d'épreuves, d'où $\mu^2 = pqr$, ce qu'il fallait démontrer.

Définition 6. — Il vaut la peine d'observer qu'on trouve le même résultat en calculant *la moyenne arithmétique des carrés des écarts absolus*. Cette moyenne arithmétique, désignons-la pour l'instant par μ^2, est, par définition,

$$\mu^2 = \frac{1}{r+1} \cdot \sum_{k=0}^{r} (k - pr)^2 \qquad \text{(IV 17)}$$

On comprend qu'il faille en prendre encore la racine carrée pour avoir la grandeur μ elle-même. La formule (IV 17) justifie dès lors le nom d'écart quadratique moyen. Or, on démontre la proposition remarquable que voici.

Théorème 2. — La grandeur définie par (IV 17) *a la même valeur que celle définie par* (IV 16). Ce théorème est un cas particulier d'une proposition générale établie plus bas (v. n° 104).

Remarque. — L'écart quadratique moyen est une des fonctions les plus importantes de la stocastique et constitue l'un des moyens d'investigation les plus puissants du statisticien.

88. — A côté de « l'écart quadratique moyen de la fréquence absolue », objet du numéro précédent, il y a lieu de considérer

l'écart quadratique moyen de la fréquence relative, μ_r. On passe du premier au second en divisant par le nombre d'épreuves, r, de même qu'on passe ainsi de l'écart absolu à l'écart relatif ou de la fréquence absolue à la fréquence relative etc. On a pour l'écart quadratique moyen de la fréquence relative, μ_r, deux théorèmes analogues à ceux du numéro précédent. En les utilisant, nous pouvons poser comme

Définition 7. — L'écart quadratique moyen de la fréquence relative est donné par

$$\mu_r = \frac{u}{r} = \sqrt{\frac{pq}{r}} \qquad \text{(IV 18)}$$

Relations entre l'écart quadratique moyen et d'autres paramètres stocastiques. En comparant l'écart quadratique moyen à l'unité d'écart u (v. n° 57), à l'écart réduit λ, à la précision de la série, $h = 1 : u$ (v. n° 77), on voit immédiatement que

$$u = \mu . \sqrt{2} \quad \text{et} \quad h = \frac{1}{\mu \cdot \sqrt{2}} = \frac{0{,}707\,1067812}{\mu} \qquad \text{(IV 19)}$$

L'écart réduit devient

$$\lambda = \frac{l}{u} = \frac{l}{\mu} \cdot 0{,}707\,106\,781 \ldots$$

Il serait facile d'augmenter le nombre de ces relations dont nous ne mentionnons ici que trois (v. n° 90 *a* et n° 92, remarque).

IV. — *L'écart probable d'une série d'épreuves répétées.*

89. — Un autre écart encore a reçu un nom qu'il faut connaître. C'est *l'écart probable*. On a critiqué ce nom en disant que n'importe quel écart a une certaine probabilité mathématique. Mais celui-ci est caractérisé par la probabilité 1/2. L'écart probable est une valeur telle que l'on a des chances égales à parier qu'elle sera dépassée par l'écart effectivement observé, ou à parier qu'elle ne le

sera pas (*). Considérons à nouveau notre série de r épreuves semblables et mettons que l'événement étudié E se réalise b fois. Si les limites $\pm c$ sont telles que le nombre de fréquence b ait autant de chances d'être compris entre $(pr - c)$ et $(pr + c)$ qu'à tomber en dehors de cet intervalle, la probabilité mathématique correspondante P_{-c}^{+c} aura la valeur $\frac{1}{2}$, et ces limites seront dites *les limites probables du nombre de fréquence*. Pour les déterminer, il faut donc résoudre par rapport à ρ l'équation

$$\Theta(\rho) + \frac{h}{\sqrt{\pi}} \cdot e^{-\rho^2} = \frac{1}{2} \qquad \text{(IV 20)}$$

Effectuons cette résolution en deux pas successifs.

a) Faisons d'abord abstraction du second terme dans le premier membre de (IV 20). On trouve (59), par la table Théta, que

$$\Theta(\rho_o) = \frac{1}{2} \text{ pour } \rho_o = 0{,}476\,936\,276\,2 \qquad \text{(IV 21)}$$

Définition 8. — Le produit $\rho_o \cdot u = \rho_o \cdot \sqrt{2pqr}$ est *l'écart absolu probable* de la série, ou *l'écart probable du nombre de fréquence*.

Les mêmes considérations que ci-dessus (v. n° 88) conduisent à poser la

Définition 9. — *L'écart relatif probable* de la série, ou *écart probable de la fréquence relative*, s'obtient en divisant l'écart absolu probable par le nombre d'épreuves r ; c'est donc $\rho_o \sqrt{\frac{2pq}{r}}$.

Définition 10. — Les *limites probables du nombre de fréquence* sont dès lors

$$p \cdot r \pm \rho_o \cdot u = p \cdot r \pm 0{,}476\,936 \cdot \sqrt{2pqr}, \qquad \text{(IV 22)}$$

(*) Mr Emile Borel, s'écartant de la terminologie usuelle, dans ses *Eléments de la théorie des probabilités*, 2e éd., Paris, 1910, p. 57, appelle « écart médian » ce que la presque unanimité des auteurs anglais, allemands, français, italiens, appelle « écart probable ».

(59) D'après le calcul refait par H. Opitz (54). Vu l'usage fréquent de cette grandeur, indiquons ici son logarithme vulgaire ; c'est

$$\log \rho_o = 9.678\,460\,356\,6 - 10.$$

et les *limites probables de la fréquence relative* sont

$$p \pm \rho_o \cdot \frac{u}{r} = p \pm 0{,}476\,936 \cdot \sqrt{\frac{2pq}{r}} .$$

Quand le nombre d'épreuves, r, devient de plus en plus grand, les premières s'élargissent indéfiniment, tandis que les secondes se resserrent toujours davantage ; c'est conforme au théorème de Jacques Bernoulli ; voir en particulier n° 81, δ.

b) Tenons compte aussi du second terme dans le premier membre de (IV 20) ; il en résultera une approximation notablement plus grande. Posons donc $\rho = \rho_o + \delta$. On démontre, v. n° 85, qu'on a très approximativement $\delta = -1 : 2u$; il faut donc retrancher 1/2 du nombre ci-dessus trouvé. On a dès lors le résultat suivant. *Limites probables du nombre de fréquence :*

$$pr \mp \frac{1}{2} \pm \rho_o \sqrt{2pqr} \qquad \text{(IV 23)}$$

Limites probables de la fréquence relative : $p \mp \frac{1}{2r} \pm \rho_o \sqrt{\frac{2pq}{r}}$.

La grandeur de « l'intervalle probable » est ainsi

$2\rho_o \sqrt{2pqr}$ — 1 pour la fluctuation, et

$2\rho_o \sqrt{2pq : r} - \frac{1}{r}$ pour l'écart relatif.

90. — *a*) En introduisant l'écart quadratique moyen $\mu = \sqrt{pqr}$ (v. n° 87), on trouve pour l'écart probable du nombre de fréquence

$$pr \mp \frac{1}{2} \pm \rho_o \cdot \sqrt{2} \cdot \mu = pr \mp \frac{1}{2} \pm 0{,}674\,489\,7 \cdot \mu \qquad \text{(IV 24)}$$

Remarquons encore, comme corollaire de la définition 8, le *Théorème* suivant : L'écart probable d'une série est directement proportionnel à l'unité d'écart absolu et inversement proportionnel à la précision de la série :

$$\text{Ecart probable} = \rho_o \cdot u = \frac{0{,}476\,936}{h} . \qquad \text{(IV 25)}$$

b) Il faut bien noter que « l'écart probable » qui vient d'être défini n'est pas l'écart *le plus* probable. En effet, c'est l'écart nul qui est le plus probable, puisque par définition l'écart zéro correspond au résultat le plus vraisemblable de la série (v. n° 76).

Problème 4. — Si l'on fait un très grand nombre de séries comprenant chacune 50 coups à pile ou face (v. n° 38), quels sont les écarts probables? *Solution.* — Il faut s'attendre à ce que les séries où pile et face se présentent 25 fois soient les plus nombreuses, car ce sont elles qui présentent l'écart nul et l'on a la quasi-certitude de voir l'écart zéro se réaliser le plus souvent. Les limites probables étant ici

$$\rho_0 \cdot \sqrt{2 \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot 50} = 5\rho_0 = 2{,}384\,68\ldots,$$

on peut parier 1 contre 1 que la moitié des séries présenteront un écart inférieur à ± 3, c'est-à-dire apparition de pile de 23 à 27 fois, et l'autre moitié des séries un écart absolu égal ou supérieur à 3, c'est-à-dire apparition de pile au moins 28 fois. Les limites probables de la fréquence relative sont ici $\pm 5\rho_0 : 50 = \pm 0{,}047\,693\,6$. L'écart quadratique moyen (de la fréquence absolue) est

$$\mu = \sqrt{pqr} = \sqrt{1/2 \cdot 1/2 \cdot 50} = \frac{10\sqrt{2}}{4} = 3{,}535\,539.$$

V. — *L'écart moyen d'une série d'épreuves répétées.*

91. — Une troisième espèce d'écart joue un rôle important dans les applications, *l'écart moyen* d'une série d'épreuves répétées r fois de suite. On entend par là *la moyenne arithmétique des écarts absolus.* C'est en quelque sorte l'analogue de l'écart quadratique moyen, mais au lieu de prendre le carré de chaque écart absolu, on en prend la première puissance. De la proposition générale établie plus bas (v. n° 104, théorème 2) et dont nous avons déjà fait usage au numéro 87, théorème 2, il suit que l'écart moyen d'une série est la somme des produits obtenus en multipliant chacun des écarts absolus possibles par sa probabilité mathématique. Or, les écarts positifs sont aussi probables que les écarts négatifs. *Si donc on prend chaque écart absolu avec son signe, on trouve zéro comme valeur moyenne.* En effet :

$$\sum_{k=0}^{r} (k - pr) \cdot \binom{r}{k} \cdot p^k \cdot q^{r-k} = 0 \qquad \text{(IV 26)}$$

On le voit immédiatement en reprenant le raisonnement du n° 87, théorème 1. Supposons qu'on exécute une $(r+1)^{\text{ième}}$ épreuve. La somme en question en sera augmentée de $p.q + q\ (-p) = 0$, c'est-à-dire qu'elle ne sera pas modifiée du tout. Donc, cette somme conserve toujours la même valeur, indépendante du nombre d'épreuves. Et comme cette valeur est nulle pour $r = 0$, la proposition se trouve établie. Aussi fait-on la convention que chaque écart doit être pris en valeur absolue, d'où la

Définition 11. — On entend par *écart moyen d'une série*, ou *valeur moyenne de l'écart*, et l'on désigne par e_a, la somme suivante :

$$e_a \equiv \sum_{k=0}^{r} abs(k - pr) \cdot \binom{r}{k} \cdot p^k \cdot q^{r-k}; \qquad \text{(IV 27)}$$

en paroles : on multiplie la valeur absolue de chacun des écarts absolus possibles par la probabilité mathématique de ce même écart. On démontre le

Théorème 1 : cette somme est égale au produit $2pqrT_o$, où T_o représente le terme maximal dans le développement de $(p+q)^r$ (voir les numéros 69 et 73). En introduisant l'expression approchée de T_o, formule (III 28), on obtient

$$e_a = \sqrt{\frac{2pqr}{\pi}} \qquad \text{(IV 28)}$$

Remarque. — On peut faire à propos de l'écart moyen d'une série la même observation que pour l'écart quadratique moyen (v. n° 87, définition 6). Le nom d'écart « moyen » se justifie par la

Définition 12. — *L'écart moyen d'une série* est la moyenne arithmétique des valeurs absolues des écarts absolus. En formule :

$$e_a = \frac{1}{r+1} \cdot \sum_{k}^{0\ldots r} abs(k - pr). \qquad \text{(IV 29)}$$

Comme au numéro 87 (théorème 2), on a la proposition remarquable que voici :

Théorème 2. — Les formules (IV 27) et (IV 29) *donnent la même valeur* e_a, bien que la première contienne des probabilités mathématiques et que la seconde soit tout-à-fait indépendante de cette notion. C'est un cas particulier de la proposition ci-dessous (v. n° 104, théorème 2). On voit aussi que c'est par la formule de Stirling que le nombre π se glisse dans l'expression (IV 28) de l'écart moyen d'une série.

92. — La formule (IV 28) donne l'écart moyen de la fréquence absolue ; et c'est en même temps sa valeur moyenne ou, d'une manière plus précise, c'est la valeur moyenne de la valeur absolue des écarts absolus. En partant des mêmes considérations que ci-dessus (v. n° 88), on est conduit à introduire *l'écart moyen de la fréquence relative*, écart que je désigne par e_r. Il est naturel de passer de l'écart moyen de la fréquence absolue à celui de la fréquence relative en divisant le premier par le nombre d'épreuves, r ; d'où la

Définition 13. — L'écart moyen de la fréquence relative, e_r, est égal au quotient obtenu en divisant l'écart moyen de la série, e_a, par le nombre d'épreuves r.

On a pour e_r des théorèmes analogues à ceux énoncés pour e_a au numéro précédent. En particulier, la grandeur e_r est en même temps *la valeur moyenne de la valeur absolue des écarts relatifs*. On peut dès lors poser

$$e_r \equiv \frac{e_a}{r} = \sqrt{\frac{2pq}{r\pi}} . \qquad \text{(IV 30)}$$

Remarque. — En introduisant l'unité d'écart absolu, $u \equiv \sqrt{2pqr}$ (v. n° 57), la précision de la série $h \equiv 1 : u$ (v. n° 77) et l'écart quadratique moyen $\mu = \sqrt{pqr}$, on a les relations suivantes :

$$e_a = \frac{u}{\sqrt{\pi}} = \frac{1}{h\sqrt{\pi}} = \mu\sqrt{\frac{2}{\pi}} = \frac{0{,}564\,1896}{h} = 0{,}797\,8936\,\mu \qquad \text{(IV 31)}$$

L'écart moyen de la fréquence absolue est donc proportionnel à l'unité d'écart absolu ; de même, l'écart moyen de la fréquence relative est proportionnel à l'écart étalon (v. n° 57, définition 7) :

$$e_r = \frac{u}{r\sqrt{\pi}} = 0{,}564\,1896 . \frac{u}{r} . \qquad \text{(IV 32)}$$

Récapitulation.

93. — En résumé, nous avons introduit, à propos de la théorie des écarts, les 15 notions réunies dans le tableau suivant. A remarquer qu'elles se groupent par deux, sauf l'écart réduit λ qui occupe une place à part. Elles forment ainsi un ensemble de sept concepts dont chacun s'applique deux fois, savoir à une fréquence absolue et à une fréquence relative.

On considère une série de r épreuves semblables dont chacune réalise nécessairement soit l'événement aléatoire E de probabilité mathématique p, soit l'événement contraire F de probabilité mathématique $p(F) = q = 1 - p$. Ces probabilités sont appelées souvent *les probabilités élémentaires*. Voici le tableau en question.

1a) La fréquence absolue ou le nombre de fréquence, k, n. Sauf indication contraire, k est le nombre d'apparition le plus probable de E, et n celui de F.

1b) La fréquence relative, $k : r$, $n : r$.

2a) L'écart absolu ou la fluctuation de la série, $l = k - p.r$ ou, si $p.r$ n'est pas un nombre entier, $l = k - En(pr - p)$ (v. n° 56).

2b) L'écart relatif de la série, $\frac{l}{r} = \frac{k}{r} - p \equiv L$.

3a) L'unité d'écart absolu, $u \equiv \sqrt{2pqr}$.

3b) L'unité d'écart relatif ou écart étalon,

$$u : r = \sqrt{2pq : r}.$$

4) *L'écart réduit* de la série, λ, important par le fait qu'il mesure à la fois l'écart absolu et l'écart relatif,

$$\lambda \equiv l : u = l : \sqrt{2pqr}.$$

5a) La précision de la série ou mesure de la précision absolue,

$$h = 1 : u = 1 : \sqrt{2pqr}.$$

5b) La mesure de la précision relative,

$$h_r = \frac{r}{u} = \sqrt{\frac{r}{2pq}}.$$

6a) *L'écart quadratique moyen* de la série, μ, ou écart quadratique moyen de la fréquence absolue,

$$\mu = \sqrt{pqr} = 0{,}707\,106\,78 . u$$

6b) L'écart quadratique moyen de la fréquence relative,

$$\mu_r \equiv \mu : r = \sqrt{\frac{pq}{r}} .$$

7a) L'écart probable de la série ou écart probable de la fréquence absolue,

$$\rho_0 . u \equiv 0{,}476\,936 \sqrt{2pqr} = 0{,}674\,489\,7 . \mu ;$$

plus exactement $\rho_0 . u - \frac{1}{2}$.

7b) L'écart probable de la fréquence relative,

$$\rho_r . \frac{u}{r} = 0{,}476\,936 . \sqrt{\frac{2pq}{r}} ; \text{ plus exactement } \rho_r . \frac{u}{r} - \frac{1}{2r} .$$

8a) L'écart moyen de la série ou écart moyen de la fréquence absolue (chaque écart étant pris en valeur absolue).

$$e_a = \sqrt{\frac{2pqr}{\pi}} = 0{,}797\,894 . \sqrt{pqr} = 0{,}7979 . \mu .$$

8b) L'écart moyen de la fréquence relative (chaque écart étant pris en valeur absolue),

$$e_r \equiv \frac{e_a}{r} = \sqrt{\frac{2pq}{r\pi}} = 0{,}797\,894 . \sqrt{\frac{pq}{r}} .$$

9) *La moyenne arithmétique des écarts est nulle, si chaque écart est pris avec son signe.* C'est pourquoi dans 8a) et 8b), chaque écart est pris en valeur absolue (voir les numéros 91 et 92).

VI. — *Limites équiprobables dans les séries d'épreuves répétées.*

94. — *Définition 14.* — Dans une suite de séries de r épreuves chacune, où chaque épreuve amène nécessairement soit l'événement

aléatoire E de probabilité mathématique p, constante pour la série envisagée, soit l'événement contraire F de probabilité mathématique $p(F) \equiv q = 1 - p$, on appelle *équiprobables* des limites qui correspondent à la même valeur de l'écart réduit λ. Elles sont caractérisées par ce fait qu'elles ont la même probabilité mathématique P_{-l}^{+l} de ne pas être dépassées par l'écart absolu l de la série ou par son écart relatif $l : r$, quel que soit le nombre d'épreuves r et la probabilité élémentaire p.

Exemple 1. — A un écart absolu $l = \pm\sqrt{pqr}$ = écart quadratique moyen μ correspond, en vertu de la définition $\lambda = l : u$ (v. n° 58), comme écart réduit :

$$\lambda = \frac{1}{2}\sqrt{2} = 0{,}707\ 107.$$

Comme $\Theta(0{,}707\ 11) = 0{,}682\ 68$, on a la probabilité mathématique

$$P = 0{,}682\ 68$$

pour que l'écart absolu d'une série ne dépasse pas son écart quadratique moyen μ. En d'autres termes, quels que soient le nombre d'épreuves r et la probabilité élémentaire p, on peut parier à peu près 68 contre 32 ou 17 contre 8, que le nombre de fréquence de l'événement étudié ne sortira pas de l'intervalle

$$< pr - \mu, pr + \mu >$$

ou, ce qui revient au même, que la fréquence relative de E restera comprise entre les limites

$$p \pm \sqrt{\frac{pq}{r}}.$$

Exemple 2. — A un écart absolu $= \pm \rho_0 u$ = écart probable de la série (v. n° 89) correspond comme écart réduit $\lambda = \rho_0 = 0{,}476\ 936$. Comme $\Theta(\rho_0) = 1/2$, on a la probabilité mathématique $P = 1/2$ pour que l'écart absolu d'une série ne dépasse pas $\pm\rho_0 u \mp \frac{1}{2}$; autrement dit, quels que soient le nombre d'épreuves r et la probabilité

élémentaire p, on peut parier 1 contre 1 que le nombre de fréquence de E ne sortira pas de l'intervalle

$$< pr - \rho_0 u + \frac{1}{2}, pr + \rho_0 u - \frac{1}{2} >$$

ou, ce qui revient au même, que la fréquence relative de l'événement E restera comprise entre les limites

$$p \pm 0,67449 \sqrt{\frac{pq}{r}} \mp \frac{1}{2r}.$$

En général, supposons que l'écart absolu l d'une série d'épreuves soit mis sous la forme $l = j.u$, où j représente un nombre réel et $u \equiv \sqrt{2pqr}$ l'unité d'écart absolu. Le nombre de fréquence le plus probable de cette série étant pr (v. n° 70), on a le

Théorème. — La probabilité mathématique P_{-l}^{+l} pour que le nombre de fréquence d'une série d'épreuves semblables soit compris dans l'intervalle $< pr - ju, pr + ju >$ est $= \Theta(j)$, c'est-à-dire que cette probabilité mathématique P_{-l}^{+l} est indépendante du nombre d'épreuves r et de la probabilité élémentaire p.

Ainsi, de ce que $\Theta(1,163) = \frac{9}{10}$ et $\Theta(2,327) = \frac{999}{1000}$, on déduit qu'il y a à parier 9 contre 1, respectivement à parier 999 contre 1, que la fréquence absolue de l'événement étudié ne sortira pas des limites

$$pr \pm 1,163.u, \text{ respectivement } pr \pm 2,327.u,$$

ou encore : que la fréquence relative de E ne s'écartera pas de sa valeur la plus probable, p, de plus de

$$1,163.\sqrt{\frac{2pq}{r}}, \text{ respectivement } 2,327.\sqrt{\frac{2pq}{r}},$$

dans un sens ou dans l'autre.

95. — Comme $u = \sqrt{2}.\mu$ et par conséquent $ju = j\sqrt{2}.\mu = k\mu$, on peut énoncer le résultat ci-dessus comme suit.

Théorème. — La probabilité mathématique que l'écart absolu d'une série d'épreuves répétées dans les mêmes conditions ne

dépasse pas $k\mu$, *où* k *représente un nombre réel et* μ *l'écart quadratique moyen de la dite série, dépend exclusivement du coefficient numérique* k. Cette probabilité mathématique ne dépend ni du nombre d'épreuves ni de la probabilité élémentaire p. Aussi a-t-on calculé des tables donnant la probabilité mathématique en question pour les valeurs successives de k. Par exemple :

La probabilité mathématique P pour que l'écart absolu ne dépasse pas, en valeur absolue,

1 fois	2 fois	3 fois	3,8 fois

l'écart quadratique moyen μ, cette probabilité est égale à

$P = 0{,}6827$	$0{,}9545$	$0{,}9973$	$0{,}9999$.

Des considérations tout analogues peuvent se faire à propos de l'écart probable de la série, $\rho_0\sqrt{2pqr}$ (v. n° 89), à propos de son écart moyen e_a (v. n° 91) et de l'unité d'écart absolu $u \equiv \sqrt{2pqr}$. On voit que *la probabilité mathématique* P *pour que la fluctuation reste inférieure à* k *fois l'unité d'écart* u *tend très rapidement vers la certitude quand* k *augmente*. Ainsi, la probabilité mathématique P pour que l'écart absolu ne dépasse pas, en valeur absolue,

1 fois	2 fois	3 fois	4 fois

l'unité d'écart absolu u, cette probabilité est égale à

$P = 0{,}842\,700\,8$	$0{,}995\,322\,3$	$0{,}999\,977\,9$	$0{,}999\,999\,984\,6$

Pratiquement, on peut donc regarder une fluctuation de $4u = 5{,}656\,85.\,\mu$, c'est-à-dire quadruple de l'unité d'écart u, comme n'étant jamais dépassée, la probabilité mathématique d'un dépassement étant de quelques cent-millionièmes seulement.

VII. — *Les formules du hasard.*

96. — Considérons de nouveau une série de r épreuves identiques. Jusqu'ici, nous n'avons envisagé que les écarts absolus, pris

avec leurs signes ou en valeur absolue (v. n° 91), et les carrés des écarts absolus (v. n° 87, définition 5). Rien n'empêche de considérer également les puissances supérieures de ces fluctuations.

Désignons, pour notre série d'épreuves,

par S_1 la somme des valeurs absolues de tous les écarts absolus,
par S_2 » » » carrés de tous les écarts absolus,
par S_3 » » » cubes des valeurs absolues de tous les écarts absolus,
par S_4 » » » 4ièmes puissances de tous les écarts absolus,
etc.

Supposons qu'on calcule la moyenne arithmétique de chacune de ces puissances ; on obtiendra

$$\frac{S_1}{r}, \quad \frac{S_2}{r}, \quad \frac{S_3}{r}, \quad \frac{S_4}{r}, \text{ etc.}$$

Or $S_1 : r$, moyenne arithmétique des valeurs absolues des écarts, n'est pas autre chose que l'écart moyen de la série (v. n° 91), donné par la formule (IV 28).

$S_2 : r$, moyenne arithmétique des carrés des écarts, est le carré de l'écart quadratique moyen de la série (v. n° 87), soit $\mu^2 = pqr$ en vertu de (IV 15). Formons le quotient de la seconde moyenne par le carré de la première. Il vient

$$\frac{S_2}{r} : \left(\frac{S_1}{r}\right)^2 = \mu^2 : e_a^{\,2} = \frac{pqr.\pi}{2pqr} = \frac{\pi}{2} \qquad \text{(IV 33)}$$

C'est la fameuse *formule du hasard*. Elle est caractérisée par ce fait que le premier membre est fourni par le hasard, tandis que le second membre est une constante connue et fixe. Par exemple, qu'on prenne 100 séries également longues ; chacune présente un écart déterminé. Qu'on prenne la moyenne arithmétique de ces 100 écarts (ce sera $e_a = S_1 : r$). Qu'on calcule ensuite les carrés de ces 100 écarts et qu'on prenne la moyenne arithmétique de ces 100 carrés. Cette deuxième moyenne divisée par le carré de la première donnera, à très peu près, le nombre $\frac{\pi}{2} = 1{,}5709$.

97. — On peut calculer les autres sommes S_ν. Il est commode d'utiliser à cet effet la grandeur h, mesure de la précision de la série (v. n° 77). On démontre alors que

$$\frac{S_1}{r} \equiv e_a = \frac{1}{h\sqrt{\pi}}; \qquad \frac{S_2}{r} \equiv \mu^2 = \frac{1}{2h^2}; \quad \text{en général}$$

$$\frac{S_m}{r} = \frac{\Gamma\left(\frac{m+1}{2}\right)}{h^m\sqrt{\pi}}, \qquad \text{(IV 34)}$$

où $\Gamma(x)$ représente la fonction Gamma introduite par Euler dans l'analyse mathématique. Cette fonction jouit entre autres des deux propriétés suivantes. 1) Pour un nombre entier positif n, on a

$$\Gamma(n) = (n-1)! \equiv 1.2.3.4.\ldots.(n-1)$$

2) Pour $x = \frac{2n+1}{2}$, on a $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ et en général

$$\Gamma\left(\frac{2n+1}{2}\right) = \frac{2n-1}{2} \cdot \frac{2n-3}{2} \cdot \frac{2n-5}{2} \ldots\ldots \frac{5}{2} \cdot \frac{3}{2} \cdot \frac{1}{2} \cdot \sqrt{\pi}$$

Il vient donc : $\frac{S_3}{r} = \frac{1}{h^3\sqrt{\pi}}$, $\frac{S_4}{r} = \frac{1.3}{4h^4}$,

$$\frac{S_5}{r} = \frac{1.2}{h^5\sqrt{\pi}}, \qquad \frac{S_6}{r} = \frac{1.3.5}{8h^6}, \text{ etc.}$$

On tire de ces relations les curieuses *formules du hasard* suivantes, dont la loi de formation est manifeste.

$$\frac{S_3}{r} : \left(\frac{S_1}{r}\right)^3 = 1!\,\pi; \quad \frac{S_4}{r} : \left(\frac{S_1}{r}\right)^4 = \frac{1.3}{2.2}\pi^2; \quad \frac{S_5}{r} : \left(\frac{S_1}{r}\right)^5 = 2!\,\pi^2;$$

$$\frac{S_6}{r} : \left(\frac{S_1}{r}\right)^6 = \frac{1.3.5}{2.2.2}\pi^3; \text{ etc.} \qquad \text{(IV 35)}$$

Ces formules singulières, dont le premier membre est fourni par le hasard, mais pas le second, « méritent tant de confiance qu'un calculateur à qui des observations sont remises et qui trouve ces égalités en défaut peut tenir pour certain qu'on a retouché et altéré les résultats immédiats de l'expérience » ([60]).

(60) J. Bertrand, *l. cit.* (49), p. 191.

VIII. — *Sur la précision d'une série d'épreuves répétées.*

98. — Supposons effectuées deux séries d'épreuves (ou d'observations). Qu'elles se rapportent au même sujet ou non, il y a de toute façon un point où elles sont comparables : c'est la longueur des intervalles équiprobables pour la fréquence des événements considérés. Lorsqu'il y a la même probabilité mathématique que, dans nos deux séries, l'écart (ou l'erreur) tombe entre les mêmes limites $-l \ldots\ldots + l$, il est naturel de dire que ces deux séries sont *également précises*. On dira au contraire que la première série est 10 fois plus précise que la seconde dans le cas suivant : la probabilité mathématique P_{-x}^{+x} pour que, dans la *première* série, l'écart (ou l'erreur) ne dépasse pas x *millimètres* est exactement la même que la probabilité mathématique pour que, dans la *seconde* série, l'écart (ou l'erreur) ne dépasse pas x *centimètres*. D'une manière générale, dire qu'une série est a fois plus précise qu'une autre signifie que la probabilité mathématique d'écarts (ou d'erreurs), compris entre certaines limites pour la première série, est la même que celle d'écarts compris entre des limites a fois plus grandes pour la seconde série. Pour arriver à une expression mathématique de la précision, envisageons 1) une première série de r épreuves dont chacune amène nécessairement soit l'événement E de probabilité mathématique $p(E) \equiv p$, soit l'événement contraire F de probabilité mathématique $p(F) \equiv q = 1 - p$; 2) une deuxième série de r' épreuves dont chacune réalise nécessairement soit l'événement aléatoire E' de probabilité mathématique p', soit l'événement contraire F' de probabilité mathématique $q' = 1 - p'$. Supposons les probabilités mathématiques p et p' constantes au cours des épreuves. L'écart étalon (v. n° 57) est par définition

$u : r = \sqrt{2pq : r}$ pour la première série et

$u' : r' = \sqrt{2p'q' : r'}$ pour la seconde.

En vertu du théorème de Jacques Bernoulli (v. n° 80), on a alors la même probabilité mathématique $\Theta(\lambda)$ [ou plus exactement celle

donnée par la formule (IV 10)] pour que l'écart relatif ne dépasse pas les limites

$$-\lambda.u : r \ldots\ldots\ldots + \lambda.u : r \text{ dans la première série,}$$
$$-\lambda.u' : r' \ldots\ldots\ldots + \lambda.u' : r' \text{ dans la seconde série.}$$

La longueur de ces intervalles équiprobables, $2\lambda u : r$ pour la première et $2\lambda u' : r'$ pour la seconde série, dépend, il est vrai, de la probabilité mathématique $\Theta(\lambda)$, mais, chose remarquable, *le rapport des dits intervalles*, $\frac{ur'}{u'r}$, *est indépendant de l'écart réduit* λ et par conséquent aussi de la probabilité mathématique $\Theta(\lambda)$. Or, plus les limites de l'écart relatif sont resserrées, plus on peut être certain que la fréquence relative se rapprochera de la probabilité élémentaire p, plus aussi la série est réputée précise. Donc, la probabilité mathématique $P_{-l}^{+l} = \Theta(\lambda)$ étant donnée et fixe, le degré de précision de la série augmente lorsque $2\lambda u : r$, longueur de l'intervalle en question, diminue, ou, ce qui revient au même, lorsque la valeur réciproque, $\frac{r}{2\lambda u}$, augmente. Comme λ est fixe (en même temps que P_{-l}^{+l} donnée par (IV 11) ou par (IV 10)), on peut dire que le degré de précision de la série augmente en même temps que $r : u = r : \sqrt{2pqr}$.

Or, cette expression n'est pas autre chose que *la mesure de la précision relative*, h_r, introduite au numéro 77, définition 2, formule (IV 8). Cette dénomination se trouve ainsi justifiée. Il s'ensuit ces deux

Théorèmes. — 1) *La mesure de la précision relative augmente proportionnellement à la racine carrée du nombre des épreuves.*

2) *Cette mesure de précision est d'autant plus grande que le produit* pq *est moindre*, par conséquent d'autant plus grande que les probabilités mathématiques p et $q = 1 - p$ sont plus différentes l'une de l'autre. Cette mesure h_r augmente donc, *ceteris paribus*, quand la probabilité élémentaire p s'éloigne de la valeur 1/2 dans un sens ou dans l'autre.

99. — Si l'on veut comparer, au point de vue purement formel, la théorie des écarts à celle des erreurs d'observation, il faut tenir compte de ce qu'une erreur d'observation doit être assimilée à un écart absolu et non à un écart relatif. Il s'ensuit que, pour ce qui concerne la mesure du degré de précision d'une série, il faut envisager non la mesure de la précision relative, comme au numéro précédent, mais *la mesure* h *de la précision absolue* ; voir à ce propos le numéro 77. Pour rendre identiques les formules des deux théories parallèles, celle des erreurs d'observation et celle des épreuves répétées, on doit définir *la mesure de la précision* par la formule (IV 5), $h = 1 : \sqrt{2pqr}$. *Elle diminue quand le nombre des épreuves augmente*, étant inversement proportionnelle à la racine carrée de ce nombre. On peut rapprocher cette circonstance du fait que la probabilité mathématique T_0 du résultat le plus probable d'une série d'épreuves répétées diminue de la même manière que h, quand le nombre d'épreuves croît ; le rapport de ces deux quantités est en effet constant, $h : T_0 = \sqrt{\pi}$.

100. — Pour comparer entre elles deux séries d'épreuves répétées, la première de précision absolue h, la seconde d'une précision absolue supérieure $H > h$, il est avantageux de recourir à la représentation graphique et de tracer la courbe d'équation

$$y = \frac{h}{\sqrt{\pi}} \cdot e^{-h^2x^2}. \qquad \text{(IV 36)}$$

Cette courbe est symétrique par rapport à l'axe des Y et a une forme qui rappelle le contour apparent d'une cloche ; on l'appelle *la courbe des probabilités* ou *courbe en cloche* ou *courbe de Gauss*, parfois *courbe binomiale* (v. fig. 1). Elle permet de lire immédiatement la probabilité mathématique d'un écart donné l. En effet, en vertu de la formule (IV 7) et du numéro 76, on a

$$y = T_x, \qquad \text{(IV 37)}$$

Donc, *première propriété* : La probabilité mathématique d'un écart absolu égal à l est égale à l'ordonnée y correspondant à l'abscisse $x = l$. *Deuxième propriété* : La probabilité mathémati-

que que l'écart absolu reste compris entre deux nombres a et b est égale à l'aire de la surface limitée par la courbe en cloche, l'axe des x et les deux ordonnées correspondant aux abscisses $x = a$ et $x = b$ (v. fig. 2). *Troisième propriété* : L'aire de la surface totale comprise entre la courbe et son asymptote est égale à 1. *Quatrième*

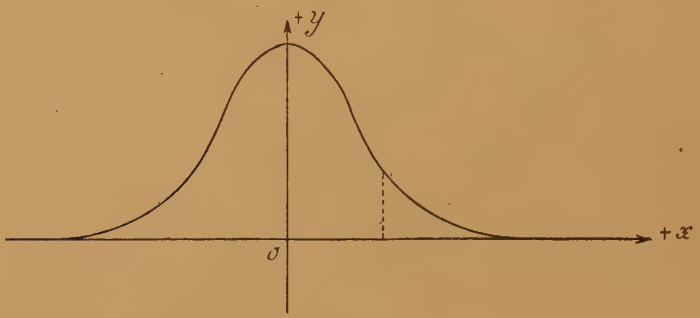


Fig. 1.

propriété : L'abscisse du point d'inflexion de la courbe correspond à la valeur de l'écart quadratique moyen μ (v. n° 87). Notons enfin ce fait remarquable que *la forme de la courbe dépend d'une seule constante*, h, appelée parfois *le module*, alors que le problème auquel elle se rapporte dépend, lui, de deux constantes, savoir : le nombre d'épreuves r et la probabilité élémentaire p.

Comparons maintenant deux courbes de probabilité ayant des modules différents, la première le module h et la seconde le module $H > h$. On démontre mathématiquement le fait suivant que l'on peut d'ailleurs constater en construisant effectivement les deux courbes : Avec la courbe de module supérieur, les petits écarts sont plus probables et les grands écarts moins probables ; plus la constante h devient grande, plus la probabilité mathématique des petits écarts augmente et plus aussi celle des grands écarts diminue (voir dans la figure 3 les deux courbes binomiales *ponctuées*). La formule (IV 19) dont on tire

$$\mu = 0{,}707\ 107 : h$$

prouve que l'écart quadratique moyen, μ, est d'autant plus faible

que h est plus grand. Il en est de même de l'écart probable de la série, $t = 0{,}476\,936 : h$, et de son écart moyen $e_a = 0{,}564\,190 : h$, voir le numéro 90 et la formule (IV 31). De tout cela, on conclut que *la précision absolue d'une série d'épreuves est d'autant plus grande que* h *est plus grand* ; par conséquent, le module h est propre à mesurer le degré de précision absolue. La dénomination de « mesure de la précision » appliquée à h se trouve ainsi justifiée.

Remarque. — Dans les applications stocastiques, notamment dans la biométrique, intervient le plus souvent une fonction dépendant de deux constantes, h et a, savoir

$$y = \frac{h}{\sqrt{\pi}} \cdot e^{-h^2(x-a)^2}. \qquad \text{(IV 38)}$$

Cette fonction atteint son maximum pour $x = a$. Mais la grandeur de ce maximum $0{,}564\,190 . h$, ainsi que la rapidité de décroissance de y de part et d'autre de l'extrême dépendent de h, comme pour (IV 36) où $a = 0$. La *forme* de la courbe dépend donc exclusivement de h ; l'effet de la constante a est de déplacer la courbe, dans son ensemble, par rapport à l'axe des y (v. fig. 2).

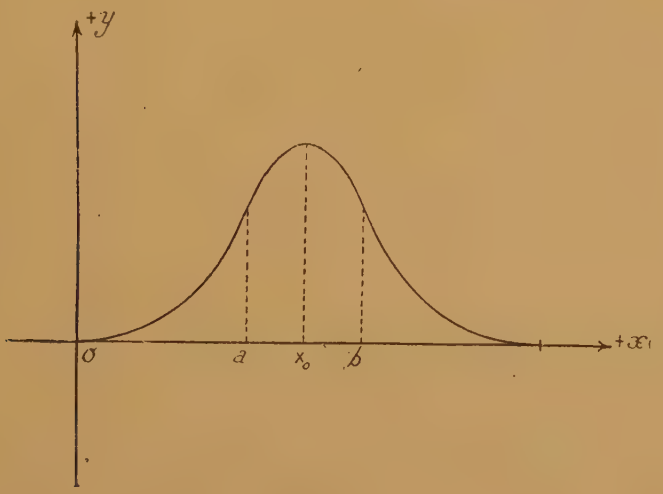


Fig. 2

IX. — *Sur la dispersion dans les séries d'épreuves répétées.*

101. — Si l'on exécute une série de r épreuves dont chacune amène nécessairement soit l'événement E soit son contraire F, de probabilités mathématiques constantes

$$p(E) \equiv p \text{ et } p(F) \equiv q = 1 - p,$$

les théorèmes ci-dessus permettent : 1) de prévoir le résultat le plus vraisemblable de la série; 2) d'en calculer la probabilité mathématique T_0 et 3) de prévoir avec quelle probabilité mathématique P_{-l}^{+l} on doit s'attendre à tel ou tel écart; rappelons que le plus important est l'écart quadratique moyen μ et que

$$P_{-\mu}^{+\mu} = 0{,}682\,68.$$

Pour pouvoir faire des considérations d'une portée plus grande, il faut exécuter, au sujet du même événement E, toute une suite de séries, chacune comportant le même nombre d'épreuves, r, et par suite la même précision h ou h_r. Supposons faites z de ces séries, et que l'on ait constaté que l'événement aléatoire E, dans la

série numéro	1,	2,	3,	.....,	z
s'est produit	k_1,	k_2,	k_3,	.....,	k_z fois.

On en déduit une série de fréquences relatives

$$\frac{k_1}{r}, \frac{k_2}{r}, \frac{k_3}{r}, \ldots\ldots, \frac{k_z}{r}. \qquad (1)$$

En vertu de la loi empirique du hasard (voir les numéros 50, 51, 54, 202), chacune est voisine de la probabilité mathématique p. Ces nombres $k_p : r$ se grouperont dès lors d'une certaine façon autour de p supposé connu, et il se posera le problème d'étudier la dispersion de ces nombres. Leur moyenne arithmétique

$$m_z \equiv \frac{k_1 + k_2 + \ldots\ldots + k_z}{r.z} \qquad (\text{IV } 39)$$

fournira vraisemblablement une valeur très approchée de la probabilité mathématique p, et si p était inconnu, on pourrait lui substituer cette moyenne m_z (v. n° 81, γ). Voici les considérations que l'on fait intervenir pour étudier la dispersion des fréquences relatives autour d'une valeur centrale, p ou m_z.

1) En vertu de la formule (IV 2), un écart positif l a la même probabilité mathématique qu'un écart négatif $-l$ de même grandeur ; voir les numéros 78 et 79. Donc, *dans les nombres de la suite (1), on doit s'attendre à trouver autant de quotients inférieurs à* p (ou à m_z) *que de quotients supérieurs à* p (ou à m_z).

2) D'après le théorème de Jacques Bernoulli (v. n° 80), il y a une probabilité mathématique déterminée $P_{-l}^{+l} = \Theta(\lambda)$ [ou plus exactement donnée par (IV 10)] pour que l'écart relatif $k_\nu : r - p$ ne dépasse pas

$$\pm\frac{l}{r} = \pm\lambda\sqrt{\frac{2pq}{r}}.$$

Il suit de là que $z.\Theta(\lambda)$ est le nombre le plus probable des quotients de la suite (1) qui tomberont à l'intérieur de cet intervalle

$$<-\frac{l}{r}, +\frac{l}{r}>.$$

Si donc on choisit arbitrairement un nombre l comme limite de l'écart absolu, ou $\frac{l}{r}$ comme limite de l'écart relatif, on en déduira l'écart réduit λ (v. n° 58) ; *et l'on devra s'attendre alors à trouver*

$$b \equiv z.\Theta(\lambda) \qquad \text{(IV 40)}$$

nombres de la suite (1) compris entre les limites $<-\frac{l}{r}, +\frac{l}{r}>$ et par conséquent $z.(1 - \Theta(\lambda)) = z - b$ nombres de la suite (1) en dehors des dites limites.

3) Mais d'autre part, plus le nombre z des séries effectuées est considérable, plus il est invraisemblable qu'on trouvera exactement le nombre $b \equiv z.\Theta(\lambda)$, en comptant combien de quotients de la suite (1) se trouvent effectivement dans le dit intervalle

$$<-\frac{l}{r}, +\frac{l}{r}>.$$

On doit au contraire s'attendre à un écart plus ou moins grand entre ce nombre effectivement constaté et le nombre le plus probable b. Le théorème de Jacques Bernoulli permet d'énoncer quelque chose sur l'étendue de cet écart-là. En effet, puisque l'unité d'écart est maintenant

$$u \equiv \sqrt{2pqr} = \sqrt{2 . \Theta(\lambda) . [1 - \Theta(\lambda)]} . \varepsilon,$$

on a la probabilité mathématique $\Theta(\xi)$ *pour que cet écart ne dépasse pas les limites*

$$\pm \xi . u = \pm \xi \sqrt{2 . \Theta(\lambda) . [1 - \Theta(\lambda)]} \, \varepsilon.$$

En d'autres termes :

Ayant fait ε séries de r épreuves chacune, puis calculé les quotients (1) et choisi l arbitrairement, on doit s'attendre tout d'abord (avec la probabilité mathématique $\Theta(\lambda)$) à ce que b de ces quotients tombent entre les limites $\pm l : r$, où b est défini par (IV 40). Mais si l'on compte ensuite combien de quotients de la suite (1) tombent effectivement dans l'intervalle $< -\frac{l}{r}, +\frac{l}{r} >$, on doit s'attendre, avec la probabilité mathématique $\Theta(\xi)$, à ce que ce nombre effectivement compté ne soit pas exactement égal à b, mais compris entre les limites

$$b - \xi . \sqrt{2b[1 - \Theta(\lambda)]} \quad \text{et} \quad b + \xi \sqrt{2b[1 - \Theta(\lambda)]}. \qquad \text{(IV 41)}$$

Par cette méthode, on peut étudier les fréquences relatives effectivement constatées et examiner si, et jusqu'à quel point, leur dispersion autour d'une valeur centrale p correspond à la répartition théorique des écarts.

102. — *Exemple.* — Pour vérifier expérimentalement les lois du hasard, Em Czuber ([61]) a utilisé les tirages de la Loterie génoise instituée officiellement en Autriche depuis deux siècles environ (v. n° 39). Ses investigations stocastiques embrassent la période de 1754 à 1886 pour les tirages effectués à Prague, et de 1771 à 1886 pour ceux effectués à Brünn. Pendant les 133 années en question,

(61) Emanuel Czuber, *Zum Gesetz der grossen Zahlen*, Prague, 1889. Voir aussi Czuber, *l. cit.* (46), p. 157-161.

le Lotto des nombres a été tiré à Prague 2854 fois sous le contrôle du Gouvernement, de sorte qu'on a là une série de $r = 2854$ épreuves exécutées avec toutes les précautions voulues, étant donnés les intérêts financiers en jeu.

1) La probabilité mathématique élémentaire $p = \frac{1}{18}$, déterminée ci-dessus (v. n° 39), est celle de la sortie d'un numéro quelconque, arbitrairement choisi d'avance parmi les 90 numéros possibles. Par exemple, on a la probabilité mathématique 1/18 de voir apparaître le numéro 1 parmi les cinq numéros qu'on tire de l'urne.

2) Dès lors, le nombre de fréquence le plus probable du numéro 1 dans une série de 2854 épreuves, puisque

$$p\,r = \frac{1}{18} \cdot 2854 = 158{,}555\ldots\ldots$$

n'est pas un nombre entier, est $En(pr + p) = 158 \equiv k$ (v. n° 70). La même chose est vraie pour tous les numéros ; pour chacun, le nombre d'apparition le plus probable est 158.

3) On peut dire alors qu'on a une suite de $z = 90$ séries comprenant chacune $r = 2854$ épreuves, et l'on peut se proposer d'examiner la dispersion, autour de la valeur centrale $p = 1/18$, de la fréquence relative du numéro 1, de celle du numéro 2 etc., enfin de celle du numéro 90.

4) *Problème 5.* — Quelle est la probabilité mathématique que le nombre de fréquence d'un numéro déterminé, par exemple du numéro 1, soit compris entre 156 et 160 ? c'est-à-dire qu'il y ait un écart absolu égal à 2 au plus, dans un sens ou dans l'autre ?

Solution. — Il faut déterminer d'abord l'écart réduit correspondant ; il est (v. n° 58)

$$\lambda = 2 : \sqrt{2 \cdot \frac{1}{18} \cdot \frac{17}{18} \cdot 2854} = 0,\ 115\,567.$$

La probabilité mathématique demandée est donc

$$P_{-2}^{+2} = \Theta(0,\ 1156) = 0,\ 129\ 25.$$

5) Le nombre le plus probable des numéros dont la fréquence

absolue ne sort pas de l'intervalle $< 156 ; 160 >$ est dès lors le plus grand nombre entier contenu dans

$$z.P = 90.0,129\ 25 = 11,63\ldots\ldots, \text{ donc } 11.$$

6) Mais il faut s'attendre à ce que, parmi les 90 numéros, il s'en trouve un peu plus ou un peu moins de 11, qui sortiront de 156 à 160 fois au cours des 2854 tirages. Cela donne lieu au

Problème 6. — Quelle est la probabilité mathématique que, sur les 90 numéros, il n'y en ait pas plus de 13 et pas moins de 9 dont le nombre de fréquence soit compris dans le dit intervalle $< 156 ; 160 >$? *Solution.* — L'écart absolu est ici ± 2. Il lui correspond, en vertu de la définition 8 (v. n° 58), un écart réduit

$$\lambda = 2 : \sqrt{2.0,129\ 25.0,870\ 75.90} = 0,444\ 36,.$$

et $\Theta(0,444\ 36) = 0,469\ 94$, mettons 0,47. On peut donc parier presque 47 contre 53 que, sur les 90 numéros, il s'en trouve au moins 9 et au plus 13 pour lesquels le nombre de fréquence est compris entre 156 et 160, limites incluses. Em. Czuber a constaté, par le dénombrement direct, qu'en réalité il s'en trouve 13 : les cinq numéros 13, 34, 40, 67 et 88 furent tirés chacun 156 fois ; les trois numéros 19, 52 et 68 chacun 157 fois ; les deux numéros 17 et 82 sortirent chacun 158 fois ; les deux numéros 15 et 90 apparurent chacun 159 fois et le numéro 58 fut tiré 160 fois.

7) L'écart probable d'une série de 2854 épreuves, avec $p = 1/18$, est (v. n° 89)

$$\rho_0.u = 0,476\ 936.\sqrt{2.\frac{1}{18}.\frac{17}{18}.2854} = 8,25\ldots\ldots, \text{ soit } \pm 8.$$

Comme le nombre de fréquence le plus probable est 158 (voir sous chiffre 2), les limites probables sont 158 ± 8, soit l'intervalle $< 150 ; 166 >$, limites comprises. On peut donc parier 1 contre 1 qu'un numéro x, arbitrairement choisi parmi les 90 possibles, sortira de 150 à 166 fois. Par le dénombrement effectif, Em. Czuber a constaté que pour 45 numéros, soit exactement pour la moitié de tous, le nombre de fréquence est en effet compris entre 150 et 166.

8) L'écart quadratique moyen de cette série est, d'après la formule $\mu = \sqrt{pqr}$ (v. n° 87),

$$\mu = \sqrt{\frac{1}{18} \cdot \frac{17}{18} \cdot 2854} = 12,\ 237.$$

Or, ayant dénombré le nombre de fréquence, k, pour chacun des 90 numéros, Em. Czuber pouvait calculer pour chacun d'eux : l'écart absolu

$$l \equiv k - En(pr + p) = k - 158, \text{ puis } l^2 = (k - 158)^2.$$

Faisant la somme de ces 90 carrés, il trouve $\Sigma l^2 = 13\,059$. Leur moyenne arithmétique est donc $13\,059 : 90 = 145,\ 1$; telle est dès lors la valeur du carré de l'écart quadratique moyen :

$$\mu^2 = 145,\ 1 \quad \text{d'où } \mu = \sqrt{145,\ 1} = 12,\ 0457.$$

C'est très près de la valeur théorique 12, 237 trouvée ci-dessus.

9) L'écart moyen de la série (v. n° 91) est, d'après les formules (IV 28) et (IV 31),

$$e_a = \sqrt{2pqr : \pi} = 0,\ 797\,893\,6 \cdot \mu = 9,\ 764.$$

L'écart moyen de la série effectivement observée est la moyenne arithmétique des valeurs absolues de tous les écarts absolus. Em. Czuber a trouvé

$$e_a \equiv \sum as \frac{k - 158}{90} = \frac{889}{90} = 9,\ 87\ldots\ldots,$$

ce qui est également en concordance remarquable avec la valeur théorique 9, 764 que nous venons de déterminer.

10) Vérifions sur cet exemple les fameuses formules du hasard (v. n° 96). La première dit $\mu^2 : e_a^2 = \frac{\pi}{2} = 1,\ 57079$; dans notre exemple, nous avons trouvé sous chiffre 8) que $\mu^2 = 145,\ 1$ et sous chiffre 9), que $e_a = 9,\ 87$; donc $e_a^2 = 97,\ 4169$. Le quotient en question donne

$$145,\ 1 : 97,\ 4169 = 1,\ 4895\ldots\ldots, \text{ soit } 1,\ 49$$

alors que la Théorie postule $\pi/2 = 1,57$. La concordance est satisfaisante, puisque l'erreur relative n'est que de 5 %.

X. — *Valeur moyenne et espérance mathématique.*

103. — *a*) Considérons une grandeur variable X qui puisse, nous le supposons, prendre les k valeurs $S_1, S_2, S_3, \ldots\ldots, S_k$ aussi facilement l'une que l'autre, ou qui ait effectivement pris les dites valeurs, chacune le même nombre m de fois au cours de $km = N$ observations, par exemple ($m = 1$) chacune seulement une fois. Il est naturel de poser comme

Définition 15. — On appelle *valeur moyenne de X* et l'on désigne par $M(X)$ la moyenne arithmétique de ces différentes valeurs,

$$M(X) \equiv \frac{S_1 + S_2 + S_3 + \ldots\ldots + S_k}{k} = \frac{1}{k} \cdot \sum_{\lambda=1}^{k} S_\lambda. \quad \text{(IV 42)}$$

Si X prenait ou avait pris l'une des valeurs plus souvent que d'autres, par exemple m_λ fois la valeur S_λ, où $\lambda = 1, 2, 3, \ldots\ldots, k$ et où

$$m_1 + m_2 + m_3 + \ldots\ldots + m_k = N$$

est le nombre total des observations, la moyenne arithmétique serait

$$M(X) \equiv \frac{m_1 S_1 + m_2 S_2 + \ldots\ldots + m_k S_k}{m_1 + m_2 + \ldots\ldots + m_k} = \frac{1}{N} \cdot \sum_{\lambda=1}^{k} m_\lambda . S_\lambda. \quad \text{(IV 43)}$$

Il n'est pas nécessaire que cette valeur moyenne coïncide avec l'une des valeurs effectivement atteintes par X ; pour des ensembles finis, ce n'est même pas le cas en général. Ainsi, la valeur moyenne du nombre d'élèves d'un établissement au cours de k années peut fort bien être fractionnaire ; et la valeur moyenne du nombre de points amenés dans une série de 6 coups où, par hasard, les 6 faces du dé se sont montrées successivement, est $3\ ^1/_2$, bien qu'aucune des faces ne porte ce nombre.

b) Supposons qu'une personne soit intéressée financièrement à un événement fortuit E de probabilité mathématique $p(E) = p$, de telle sorte que, si E se réalise, elle obtienne un gain de montant a.

On appelle habituellement, d'un terme assez mal choisi mais consacré par l'usage, le produit $p.a$ « l'espérance mathématique » de la personne en question. S'il s'agissait d'une perte d'argent du montant de a, on donnerait à a le signe —; l'espérance mathématique serait alors dite négative et $= -p.a$.

Définition 16. — On appelle « *espérance mathématique* d'une personne relativement à un gain aléatoire » le produit du gain possible par la probabilité mathématique de le réaliser. On démontre aisément que la notion d'espérance mathématique jouit de la propriété additive; cela donne le

Théorème 1. — L'espérance mathématique totale qui s'attache à plusieurs événements fortuits est égale à la somme des espérances mathématiques correspondant à chacun d'eux.

104. — Supposons qu'une grandeur X dépende du hasard comme suit : X prend la valeur S_1 ou S_2 ou S_3 etc. ou enfin S_k, suivant que l'événement fortuit E_1 ou E_2 ou E_3 etc. ou enfin E_k se réalise, événements dont les probabilités mathématiques, réduites au même dénominateur N, sont respectivement

$$p_1 = \frac{m_1}{N}, \; p_2 = \frac{m_2}{N}, \; p_3 = \frac{m_3}{N}, \; \ldots\ldots, \; p_k = \frac{m_k}{N}. \qquad (2)$$

On peut envisager X comme variable *casuelle* (voir la définition 17) susceptible de prendre les valeurs S_n auxquelles sont attachées les dites probabilités mathématiques p_n ($n = 1, 2, 3, \ldots\ldots, k$). Quelle est, dans ces conditions, la probabilité mathématique de S_1 par exemple ? Evidemment $m_1 : N$, puisque par hypothèse X ne prend la valeur S_1 que si E_1 se réalise et que $p(E_1) = p_1 = m_1 : N$; de même, la probabilité mathématique de S_n est $m_n : N$ pour toutes les valeurs de $n = 1, 2, 3, \ldots\ldots, k$.

Supposons que les k événements E_n forment une disjonction complète (v. n° 31, définition 7), c'est-à-dire que la réalisation de l'un d'entre eux exclue celle de tous les autres et que l'un doive nécessairement arriver. Dans ce cas, $p_1 + p_2 + p_3 + \ldots\ldots + p_k = 1$.

Cela posé, formons les k produits $S_n \cdot p_n$ et additionnons-les ; il vient :

$$p_1 S_1 + p_2 S_2 + \ldots + p_k S_k$$
$$= \frac{1}{N}(m_1 S_1 + m_2 S_2 + \ldots + m_k S_k) = M(X) \quad \text{(IV 44)}$$

en tenant compte de (IV 43). Cette égalité contient un résultat inattendu que l'on peut formuler comme suit.

Théorème 2. — La valeur moyenne d'une quantité casuelle X *est égale à la somme algébrique des produits obtenus en multipliant chacune des valeurs que* X *est susceptible de prendre, par la probabilité mathématique de la dite valeur.*

Définition 17. — Par « quantité *casuelle* », nous entendons une variable, continue ou non, dont la valeur dépend du hasard ou d'événements aléatoires.

En tenant compte de la définition 16 et du théorème 1 du numéro 103, on a le

Théorème 3. — La valeur moyenne d'une variable casuelle est égale à l'espérance mathématique totale qui s'y rattache.

Remarque 1. — Dans les applications, les probabilités mathématiques p_n ne sont pas toujours connues. On est souvent obligé de leur substituer les fréquences relatives $k_n : r$; et la moyenne arithmétique des valeurs effectivement atteintes par la variable considérée, savoir

$$\frac{1}{r}(k_1 S_1 + k_2 S_2 + \ldots + k_\nu S_\nu),$$

est une « valeur moyenne *empirique* de la grandeur X ». La valeur moyenne théorique $M(X)$, donnée par (IV 44), n'est calculable que si les probabilités mathématiques p_n sont toutes assignables, et l'on a dans ce cas

$$M(X) = \sum_n p_n \cdot S_n \cdot \quad \text{(IV 44)}$$

Remarque 2. — La définition 15 du numéro 103 s'étend au cas où la variable X est continue. Elle est alors susceptible de prendre une infinité de valeurs. On maintient dans ce cas les théorèmes 1 à 3, en remplaçant les sommes par des intégrales définies appropriées.

Exemple. — On peut concrétiser en quelque sorte par le schéma de l'urne les définitions abstraites qui précèdent. Envisageons une urne contenant N boules identiques à tous égards et ne différant les unes des autres que par un nombre inscrit sur chacune d'elles. Supposons que m_1 boules portent le nombre x_1, que m_2 autres boules portent le nombre x_2 et ainsi de suite, enfin que m_k autres boules portent le nombre x_k, où $m_1 + m_2 + m_3 + \ldots + m_k = N$.

En tirant au hasard une boule de cette urne, on en sortira un nombre X susceptible de prendre les k valeurs différentes x_1 ou x_2 ou x_3 etc. ou enfin x_k, chacune avec une probabilité mathématique déterminée indiquée dans la série (2) ci-dessus. La moyenne arithmétique des nombres inscrits sur les boules est, par définition, leur somme divisée par leur nombre ; cela donne, formule (IV 43),

$$M(X) = (m_1 x_1 + m_2 x_2 + m_3 x_3 + \ldots + m_k x_k) : N$$

$$= x_1 \cdot \frac{m_1}{N} + x_2 \cdot \frac{m_2}{N} + x_3 \cdot \frac{m_3}{N} + \ldots + x_k \cdot \frac{m_k}{N}$$

$$= x_1 \cdot p_1 + x_2 \cdot p_2 + x_3 \cdot p_3 + \ldots + x_k \cdot p_k.$$

C'est bien ce que dit le théorème 2.

CHAPITRE V

LES DIVERSES INTERPRÉTATIONS DU CALCUL DES PROBABILITÉS

105. — Dans l'introduction historique, nous avons touché aux difficultés que soulève la Théorie des chances aux points de vue logique et philosophique (voir les numéros 1 et 2). Analysons maintenant les diverses interprétations que l'on a données ou que l'on peut donner des propositions de cette théorie.

I. — *L'interprétation psychologique.*

Elle consiste à dire que tout théorème du calcul des probabilités énonce quelque chose sur un état d'âme réellement existant, sur le degré de certitude avec laquelle, à un moment donné, une personne croit ou attend quelque chose. Ce sont les doctrines de Fechner sur la mesure des grandeurs psychiques qui ont, sinon suscité cette interprétation psychologique, du moins beaucoup contribué à la répandre. Il est certain que les degrés les plus divers de certitude subjective peuvent avoir lieu en nous. A l'un des extrêmes la valeur zéro, lorsqu'une supposition ou une hypothèse ne nous paraît pas plus vraisemblable ni plus admissible que l'hypothèse contraire ; à l'autre extrême une certitude absolue, par exemple la conviction avec laquelle nous affirmons notre propre existence ou attendons le prochain lever du soleil. Entre ces deux extrêmes, une suite de degrés et de nuances innombrables dont le calcul des probabilités permet de déduire des conclusions ou de tirer des conséquences pratiques.

106. — *Critique de cette interprétation.* — On peut faire deux sortes d'objections à cette interprétation psychologique de la théorie des chances. *a*) Tout d'abord, la certitude subjective n'est pas mesurable, pas plus que les grandeurs psychiques intensives en général. Il existe bien des différences de degré et il est possible de les comparer, mais ils se dérobent à une mesure numérique précise, parce qu'ils ne sont pas composés d'éléments tout-à-fait homogènes. Nous pouvons bien dire : L'événement A nous paraît un peu plus probable ou beaucoup plus probable ou infiniment plus vraisemblable que l'événement B ; mais la phrase « A nous paraît $3\,{}^{1}/_{2}$ fois plus vraisemblable que B » n'a pas de sens précis assignable, pas plus que la proposition : « Ce tableau du peintre A nous paraît $3\,{}^{1}/_{2}$ fois plus beau que le tableau de B ». C'est un assemblage de mots vide de sens. Il est illusoire de vouloir chercher des comparaisons numériques dans ce domaine. Prenons un exemple : Avec quelle probabilité attends-je qu'une guerre n'éclate entre l'Angleterre et la Russie au cours de l'année prochaine ? Pour répondre à une telle question, il n'y aurait pas du tout à envisager les circonstances politiques réelles ; si l'on adoptait l'interprétation psychologique, il s'agirait uniquement de mon degré de conviction ou de présomption en ce moment et sans aucun égard aux causes de cet état d'âme, de cette crainte ou de cette espérance. Pour un degré de conviction ainsi entendu, il n'y a pas d'unité de mesure, pas plus qu'il n'y en a pour l'intensité de la foi ou d'un sentiment.

b) Même en laissant de côté la question de la possibilité théorique de mesurer numériquement des grandeurs comme la crainte et l'espérance, on peut démontrer facilement que les théorèmes du calcul des probabilités ne doivent pas être interprétés de cette manière. Si l'on admettait une définition purement psychologique de la probabilité mathématique, un même événement devrait avoir des probabilités mathématiques différentes suivant les personnes ; et ce n'est pas le cas. Pour déterminer une probabilité mathématique, on ne tient aucun compte du sentiment ni d'un certain nombre de facteurs qui exercent une grande influence sur notre état psychique individuel. Ainsi, la probabilité mathématique d'amener le sonnez au jeu de trictrac est $= 1/36$ (v. n° 23, problème 2) et cela indépendamment de toute circonstance individuelle. Si nous

faisions remarquer que, pour beaucoup de joueurs, l'apparition du double six entraîne un grand gain pécuniaire et que ce fait augmente leur espoir de voir arriver ce point plutôt qu'un autre, on nous répondrait avec raison que, dans les susdits théorèmes, il ne s'agit nullement de mesurer des états psychiques. L'objet du calcul des probabilités, ajouterait-on sans doute, c'est seulement le degré de certitude avec lequel il faut s'attendre « raisonnablement » à l'arrivée de tel ou tel événement.

Nous pouvons conclure de ce résultat négatif que les propositions du calcul des probabilités n'énoncent en tout cas rien sur l'état psychique qui, à un moment donné, a lieu dans un individu donné.

II. — *L'interprétation pratique.*

107. — Cette deuxième interprétation envisage les rapports du calcul des probabilités avec ce qui est pratiquement raisonnable ; elle peut se résumer comme suit : Les propositions du calcul des probabilités disent quelque chose sur le degré de certitude avec lequel il faut attendre un événement aléatoire *raisonnablement*, c'est-à-dire abstraction faite de tout sentiment et de tout facteur psychologique.

Cette interprétation n'est pas plus soutenable que la première, du moins dans le sens habituel du mot « raisonnable ». N'est-il pas raisonnable, *ceteris paribus*, de nous attendre davantage à un événement aléatoire lorsqu'il est indésirable que lorsque nous le souhaitons ? D'ailleurs, à ce point de vue, beaucoup de facteurs mériteraient d'être considérés ; par exemple les mesures que nous pourrions prendre si l'événement redouté se réalisait, ou pour empêcher son arrivée etc. On le voit : le mot « raisonnablement », introduit pour interpréter les propositions de la théorie des chances, ne doit pas être pris dans son sens éthique ordinaire. Des règles de conduite sont quelque chose de trop personnel et arbitraire pour qu'on puisse songer à en faire l'objet du calcul des probabilités. Celui-ci n'aurait pas la portée objective et l'importance qu'il revêt en fait, s'il se rapportait à des grandeurs purement sub-

jectives La validité générale des théorèmes en question, le fait que, pour déterminer la probabilité mathématique, on ne tient aucun compte de tous ces facteurs qui pourtant ne sont point négligeables pour nos actions journalières, tout cela exclut cette seconde interprétation au même degré que l'interprétation psychologique.

Remarque. — Les théories que nous venons d'exposer et de critiquer sont répandues surtout chez les logiciens anglais. Par exemple J. St. Mill (62) écrit dans sa *Logique* que la probabilité n'est point une propriété des choses en elles-mêmes, mais simplement un nom pour l'intensité avec laquelle nous les attendons. De même De Morgan (63) : Par degré de probabilité, nous entendons ou devrions entendre en réalité le degré de croyance (« the degree of belief ») ; et à leur suite, de nombreux auteurs anglais et allemands expriment des idées analogues.

III. — *L'interprétation logique.*

108. — A ce troisième point de vue, les valeurs numériques figurant dans les énoncés du calcul des probabilités ne mesurent ni un état psychique réellement existant, ni quelque chose d'exigible en vertu de règles pratiques ; ces valeurs mesurent un rapport logique, et voici comment. Quand on se place au point de vue de la théorie classique, toute recherche de probabilités mathématiques peut se ramener, au point de vue logique, à la considération de cas également possibles. L'objet de la théorie des chances est précisément de les établir, d'en mesurer l'ensemble et de le comparer au sous-ensemble de ceux qui amènent l'événement étudié. Si l'on envisage ainsi le côté logique du problème, on arrive facilement au *principe de l'absence de raison justifiante* (v. n° 18) : des cas doivent être envisagés comme *équipossibles* si, d'après l'état de nos connaissances, il n'y a aucune raison de considérer l'un comme plus vraisemblable que l'autre.

(62) John Stuart Mill (1806-1873). *System of Logic*, London, 1843 ; 8e éd., 1881, t. II, p. 67.

(63) De Morgan. *Formal Logic*, p. 172.

A première vue, cette interprétation semble donner pleine satisfaction. En effet, les facteurs susmentionnés qui peuvent exercer une si grande influence sur l'état d'âme de celui qui espère ou qui redoute la réalisation de l'événement fortuit, tout le monde s'accorde à dire qu'ils n'entrent point en ligne de compte pour la logique formelle. Quand on lance les dés, même si l'apparition du double six procurait un grand gain au joueur, tout le monde admet que, du point de vue abstrait et purement logique, le joueur n'a aucune raison de s'attendre à voir arriver le double six plutôt que le double deux ou le double trois ou toute autre combinaison équipossible de points. Donc, en adoptant l'interprétation purement logique, les facteurs psychologiques et éthiques sont éliminés, comme cela doit être.

Première objection à cette interprétation purement logique. — Le principe contient encore une très grande part de subjectivité, puisqu'il dit que la classification en cas équipossibles dépend de l'état de nos connaissances. Il peut donc arriver que deux cas, a et b, soient également probables aux yeux de X, tandis que Y qui possède des connaissances plus étendues déclare ces deux cas inégalement possibles, parce qu'il connaît une raison péremptoire qui donne plus de chances à a qu'à b. Ceci paraît en contradiction avec la validité générale et la précision que tout le monde s'accorde à attribuer aux théorèmes du calcul des probabilités Un logicien anglais écrit : « Si A et C représentent deux choses absolument inconnues, nous n'avons aucune raison de croire que A soit C plutôt que de croire que A ne soit pas C; la probabilité que A soit C est donc $= \frac{1}{2}$ ». On ne peut s'empêcher de présumer que la probabilité mathématique dont il est question dans la théorie des jeux de hasard a tout de même une base plus solide que cette ignorance complète et absolue.

Réfutation de cette objection. — Dans les jeux de hasard, les conditions mécaniques et physiques qui amènent tel ou tel résultat sont à ce point compliquées que même les plus grands savants se trouvent à leur égard sur un pied d'égale ignorance ; dès lors, en

regard des diverses possibilités, la présomption logique est aussi approximativement la même pour tout le monde.

109. — La condition énoncée au numéro précédent est certainement nécessaire : pour qu'il y ait égale possibilité de cas, aucun ne doit paraître plus probable qu'un autre, d'après l'état de nos connaissances. Cette condition nécessaire est-elle aussi suffisante ? En y réfléchissant, on verra que non, et c'est *la deuxième objection* que l'on peut faire à l'interprétation logique. Nos connaissances relatives à un événement peuvent être incomplètes de bien des façons et à des degrés très différents. *Exemple 1.* — Supposons qu'on sache d'une urne $U(a, a ; 2a)$ qu'elle contient autant de boules blanches que de noires (v. n° 33, définition 8) ; la probabilité mathématique d'en sortir une boule blanche est $p = 1/2$ (v. n° 23, problème 3). Supposons maintenant qu'on sache seulement que l'urne contient un certain nombre de boules, mais qu'on ne sache absolument rien de leur couleur, blanche ou verte ou rouge ou noire etc. En vertu du principe de la raison manquante, on devra poser la probabilité mathématique d'en sortir une boule blanche $p = 1/2$, car la boule tirée au hasard pourra tout aussi bien être blanche que non-blanche. Or, seule la première supposition répond à ce que l'on entend généralement en appliquant le calcul des probabilités, parce qu'il y a des raisons objectives de poser $p = 1/2$. Dans le deuxième cas, cette base objective fait défaut et l'état intellectuel est tout différent.

Exemple 2. — Considérons un événement qui puisse se produire sur n'importe quel point du Globe terrestre, par exemple la chute d'un bolide. Quelle est la probabilité mathématique qu'il tombe en Europe ? On pourrait raisonner comme suit : La surface terrestre comprend 5 continents et 4 océans ; d'après l'état de mes connaissances, il n'y a aucune raison d'admettre que l'une de ces 9 parties sera touchée plutôt qu'une autre ; la probabilité d'être touchée est donc 1/9 pour l'Europe (v. n° 23). Mais on pourrait aussi subdiviser la surface du Globe en n parties, par exemple suivant les différents Etats. En vertu du même principe, la probabilité pour cha-

que Etat d'être touché serait $1/n$; puis, appliquant le théorème d'addition (v. n° 32), la probabilité pour l'Europe étant égale à la somme des probabilités relatives aux différents Etats européens, on trouverait un résultat différent de 1/9. Et l'on arriverait à d'autres résultats, si l'on partageait la surface du Globe encore autrement. On le voit : dès que l'on adopte le principe de l'absence de raison justifiante, c'est l'arbitraire presque absolu dans l'établissement de cas également possibles ; voir le numéro 111.

110. — *Exemple 3.* — Quelle est la probabilité qu'un astre contienne un des éléments chimiques terrestres, par exemple que l'étoile Sirius contienne du fer? Mettons-nous à la place d'une personne ignorant l'existence de l'analyse spectrale, ou mieux encore : transportons-nous à l'époque antérieure à cette découverte. N'ayant pas plus de raisons de supposer que Sirius contient du fer que de supposer le contraire, on posera la probabilité mathématique de chacune de ces deux hypothèses égale à 1/2, en vertu du principe de l'absence de raison justifiante.

Même réponse à chacune des questions : Sirius contient-il de l'or? de l'argent? de l'hydrogène? du carbone ? etc. De plus, ces probabilités, dont chacune $= 1/2$, sont indépendantes et le théorème de multiplication est applicable (v. n° 34). De cette façon, on arriverait au résultat qu'il y a une probabilité mathématique excessivement petite, savoir

$$1 : 2^{96} = 1{,}262\,187 \ldots\ldots : 10^{28} = 0{,}000 \ldots\ldots 0\,126\,219$$

(27 zéros après la virgule), de l'ordre du dix-mille-quatrillionième, donc une probabilité pratiquement nulle, pour que les 96 éléments chimiques actuellement connus manquent tous sur Sirius et par conséquent la quasi-certitude que l'un au moins des dits éléments s'y trouve. Mais d'autre part, on peut d'emblée poser l'alternative comme suit : ou bien Sirius contient l'un des éléments chimiques actuellement connus, ou bien il n'en contient aucun. En vertu du même principe philosophique, on devrait considérer ces deux hypothèses comme également vraisemblables et dès lors poser leurs probabilités mathématiques égales à 1/2 ; ainsi, la probabi-

lité qui tout à l'heure était presque = 1 se trouve maintenant réduite à 1/2. On pourrait imaginer des alternatives conduisant à d'autres réponses encore. Le seul résultat certain est que toute cette décomposition en cas équipossibles semble inapplicable, parceque manquant de base objective et dès lors trop arbitraire.

111. — D'une manière générale, il peut arriver ceci : d'une part, qu'il n'y ait point de raison de considérer comme inégalement probables un certain nombre de cas, par exemple les six cas a, b, c, d, e, f; chacun aurait donc une probabilité mathématique $= 1/6$; d'autre part, en faisant des quatre premiers cas, a, b, c, d, la possibilité A et des deux derniers, e et f, la possibilité B, qu'il n'y ait pas non plus de raison de considérer A comme plus probable que B ; chacune de ces deux possibilités aurait donc une probabilité mathématique $= 1/2$. Mais le théorème d'addition (voir le n° 32) donne pour A la probabilité mathématique $4/6 = 2/3$ et pour B la probabilité $1/3$. On trouverait donc pour un même événement A des probabilités mathématiques différentes suivant la méthode employée. La même chose se produirait dans des cas plus généraux, et cela ferait perdre à la théorie des chances presque toute sa valeur.

On doit en conclure que, dans l'interprétation ci-dessus donnée, il manque un point essentiel : l'établissement des cas équipossibles doit être exempt d'arbitraire ; le principe de l'absence de raison justifiante exprime une condition nécessaire, mais non suffisante. On ne doit poser des cas également vraisemblables que lorsqu'on peut se baser sur des circonstances objectives *obligeant* à les estimer comme tels et permettant de déclarer fausse toute autre estimation. Sinon, la représentation du rapport logique par des nombres n'a pas grande valeur.

Quelle peut bien être la nature de ces circonstances objectives ? Voilà ce qu'il s'agira de préciser. L'exemple 2 du numéro 109 permet déjà de présumer que ce seront des rapports de grandeurs ou de certaines étendues. Ainsi, s'agissant de déterminer la probabilité mathématique pour un pays d'être touché par un bolide, si l'on pose cette probabilité proportionnelle à la surface du pays, tout arbitraire se trouve exclu ; en effet, quelle que soit la manière de

partager la surface terrestre en parties, la probabilité en question sera la même pour deux pays, s'ils ont même surface. Trouve-t-on aussi dans la théorie des jeux de hasard des rapports de surfaces ou de volumes ou d'autres grandeurs mesurables? Avant d'aborder cette question, discutée très en détail par Jean de Kries ([64]), examinons encore une autre interprétation des théorèmes du calcul des probabilités.

IV. — *L'interprétation empiriste.*

112. — Cette quatrième interprétation consiste à voir dans les théorèmes du calcul des probabilités le résumé de certaines connaissances relatives aux nombres de fréquence d'événements passés. Les théorèmes en question exprimeraient ce qui est réellement arrivé dans des cas déjà observés. Notre présomption ou conjecture relative à un événement déterminé E est naturellement influencée par ce que nous savons déjà de cas semblables; or, dire que la probabilité mathématique de E est $= f : s$ ne signifierait pas autre chose que ceci : on a déjà observé s cas semblables dans lesquels l'événement E s'est réalisé f fois (ou il s'est réalisé mf fois dans ms cas observés), les autres $s - f$ (ou $ms - mf$) cas observés ne lui ont pas été favorables.

Remarquons que des évaluations numériques entrent ici en jeu pour la détermination de notre expectative et que, sur ce point important, l'interprétation empiriste est en accord avec le principe caractéristique du calcul des probabilités. Remarquons ensuite que la base de cette interprétation est *le raisonnement par analogie*. Il convient dès lors de l'analyser de plus près. Lorsqu'on a observé x cas tout-à-fait semblables qui ont abouti chaque fois à un événement déterminé E, on s'attend, par analogie, à ce qu'un nouveau cas semblable aboutisse à son tour au résultat E. Il suit de là : 1° que la sûreté d'un raisonnement analogique augmente avec le nombre des cas déjà observés ; 2° qu'elle est fonction du degré

(64) Johannes von Kries, *Die Prinzipien der Wahrscheinlichkeitsrechnung*, Eine logische Untersuchung. Freiburg i. B., 1886.

plus ou moins grand de similitude que les cas observés présentaient entre eux et avec celui actuellement envisagé, ainsi que du caractère de cette similitude. Mais on ne peut pas expliciter mathématiquement cette fonction (v. n° 120, *b*).

Définition 1. — En général, l'événement étudié se réalise seulement dans un certain pour cent des cas observés. Ce sera l'*analogie partielle*. L'autre éventualité, où E s'est réalisé dans tous les cas observés, sera dite *analogie totale*.

Définition 2. — L'analogie partielle est toujours caractérisée par une fraction proprement dite que j'appellerai *le nombre analogique* (65). Il indique précisément quelle fraction de tous les cas observés jusqu'ici ont effectivement abouti au résultat E. Si le nombre analogique augmente, c'est que la fréquence relative de E devient de plus en plus grande. Quand le nombre analogique est $= 0$ ou $= 1$, l'analogie partielle s'est transformée en analogie totale. En adoptant cette terminologie, on peut dire que l'interprétation empiriste consiste à envisager les probabilités mathématiques comme des nombres analogiques (v. n° 138 et n° 202).

113 — *Objections contre l'interprétation empiriste.* — On peut en soulever plusieurs dont voici les principales. 1) L'expérience postulée par cette interprétation n'existe pas toujours. Par exemple, on applique avec la plus entière confiance les règles du calcul des probabilités à des jeux de hasard tout-à-fait nouveaux ou à des combinaisons que l'on n'a jamais encore essayées. L'interprétation empiriste est donc incapable d'expliquer les déterminations *a priori* de la probabilité mathématique.

2) D'autres fois, il est impossible de faire l'expérience voulue, à cause du nombre énorme d'épreuves qu'il faudrait exécuter. Qu'on songe au théorème général : « Le point amené en lançant un dé à jouer est indépendant du résultat de tous les autres jets du même

(65) On verra plus loin qu'il convient de l'appeler « probabilité empirique ». Voir à ce sujet l'article de Mr Corrado Gini, *Che cos' è la probabilità ?*, dans « Scientia » II, vol. III, Milan, 1908, p. 337-361.

dé » et à l'ensemble illimité des conséquences de ce théorème, par exemple au corollaire suivant : En lançant 6 fois de suite un dé idéal, il est tout aussi probable d'amener successivement 1, 2, 3, 4, 5, 6, que d'amener 6, 5, 4, 3; 2, 1 ou n'importe laquelle des 720 permutations possibles (v. n° 34, remarque). Personne ne conteste l'exactitude de cette proposition, mais elle n'est pas interprétable dans le sens empiriste, car les expériences dont elle devrait représenter le résumé sont trop nombreuses pour jamais pouvoir être effectuées. On ne possède que dans un nombre restreint de domaines les résultats d'expériences un peu étendues sur la constance de la fréquence relative ; néanmoins, on étend la stocastique à des domaines toujours nouveaux.

3) L'interprétation empiriste est impuissante à expliquer les différences constatées dans le résultat d'observations effectives faites pour déterminer une même probabilité mathématique.

4) La valeur logique des probabilités mathématiques est tout autre que celle des nombres analogiques. Ainsi : α) dans le calcul des probabilités, le nombre 1 exprime la certitude pleine et entière ; β) toute probabilité mathématique exprime une présomption bien déterminée, une expectative qui ne varie pas d'un cas à l'autre pour les cas rentrant dans une même catégorie ; c'est même cela qui fait l'importance de la théorie des chances. Les nombres analogiques n'ont pas ce sens. Le nombre 1 par exemple exprime l'analogie totale (v. n° 112) ; ce n'est point la certitude. Et le degré de certitude basé sur ces analogies totales ou partielles est variable d'un cas à l'autre, bien que les cas rentrent dans une même catégorie. Conclusion : *les théorèmes du calcul des probabilités contiennent plus qu'une simple indication historique sur des événements passés.*

Remarque 1. — La seule chose acceptable dans l'idée empiriste est ce qui a trait au développement historique de la théorie. On peut dire que si l'on n'avait pas remarqué la constance presqu'absolue de la fréquence relative de certains événements, le calcul des probabilités n'aurait pas été inventé. Mais du fait que l'expérience donne lieu à une notion, il ne suit pas encore que l'expérience en soit nécessairement la source philosophique.

Remarque 2. — A propos du raisonnement par analogie, faisons ressortir certaines difficultés que présente l'interprétation des théorèmes du calcul des probabilités lorsqu'on adopte le principe de l'absence de raison justifiante. Quand on ne sait absolument rien sur les causes d'un événement E, on est porté à appliquer le raisonnement par analogie ; on considérera E comme plus probable que F, si E s'est déjà produit souvent et F rarement ou pas du tout. En vertu de l'analogie, on devrait donc considérer l'arrivée du point 6 par exemple comme plus probable que celle de n'importe quel autre point, si 6 avait été amené déjà plusieurs fois par le même dé ; les suites telles que 6, 6, 6, 6 ou 2, 2, 2, 2, etc., devraient être considérées comme plus probables que 4, 6, 3, 1 ou 5, 3, 2, 4, etc. C'est aussi ce que l'on pensera si l'on a amené le même point une douzaine de fois de suite : on supposera que le dé utilisé est pipé ou qu'il y a des *circonstances favorisantes*. Or, dans la théorie classique on n'applique pas le raisonnement par analogie ; tous les cas y sont envisagés comme équipossibles. On peut déduire de là que le principe de l'absence de raison justifiante est une base logique insuffisante.

V. — *L'interprétation inductive.*

114. — Commençons par quelques considérations préliminaires sur l'induction. Toute connaissance à priori est *atemporelle*, en dehors du temps ; une fois acquise, la connaissance *apriorique* est valable pour toujours ; les jugements aprioriques ont le caractère de la certitude. Ainsi, quiconque a reconnu que deux grandeurs égales à la même troisième sont égales entre elles a acquis une connaissance certaine et « atemporelle ». Il en est tout autrement de la connaissance empirique ; ici, le moment présent forme une barrière : ce qui se trouve en deçà peut être connu par la perception ; ce qui est au delà se trouve dans l'avenir et dès lors est inaccessible à la connaissance directe. L'intelligence cherche à franchir cette barrière et à connaître ce qui n'a pas encore été perçu. Pour le penseur, cette tendance puissante cherchant à pénétrer l'avenir pose un problème : quelle est la nature de la connais-

sance se rattachant à des choses non encore vécues ni perçues par personne ? C'est le problème de *l'induction*, car c'est ainsi que l'on appelle tout jugement qui, basé sur quelque chose de perçu ou de connu, conclut à ce qui ne l'a jamais encore été.

Il faut exclure de ce problème la méthode que l'on appelle souvent en mathématiques « le procédé de l'induction complète » ; c'est cette méthode de démonstration que Descartes a mise en honneur et qui consiste à « conclure de n à $n + 1$ ». Elle s'applique seulement aux choses abstraites et n'entre pas en question dans les investigations de ce chapitre.

On peut se demander ce que l'induction vient faire dans le calcul des probabilités. Laplace a répondu à cette question d'une manière particulièrement nette dans son magistral *Essai sur les probabilités* (v. n° 11). Il écrit entre autres (p. 1) « On peut dire, à parler en rigueur, que presque toutes nos connaissances ne sont que probables et dans le petit nombre des choses que nous pouvons savoir avec certitude, dans les sciences mathématiques elles-mêmes, les moyens de parvenir à la vérité sont fondés sur les probabilités, en sorte que le système entier des connaissances humaines se rattache à la théorie exposée dans cet *Essai*... ». Plus loin (p. 168), parlant des divers moyens d'approcher de la certitude, il écrit : « *L'induction*, *l'analogie*, des hypothèses fondées sur les faits et rectifiées sans cesse par de nouvelles observations, tels sont les principaux moyens de parvenir à la vérité ».

Tous les essais de légitimer l'induction ont prouvé qu'il n'est pas possible de ramener philosophiquement l'acte intellectuel de l'induction à quelque chose d'indépendant de toute empirie. Il faut dès lors considérer l'induction comme un élément donné de la pensée humaine, comme une activité intellectuelle indépendante, comme un procédé nouveau et *sui generis* de connaître la réalité.

115. — *L'interprétation inductive* du calcul des probabilités consiste à dire que ses propositions expriment certaines lois de la nature ou certaines régularités trouvées empiriquement dans le cours des événements. La fréquence relative $f : s$ d'un événement E dans une série de s cas déjà observés (v. n° 52, n° 53) ne doit pas être considérée sans autre comme sa probabilité mathématique $p(E)$.

Cependant, le fait que le dit événement s'est effectivement réalisé tant et tant de fois dans une nombreuse série permet de tirer la conclusion qu'il en sera toujours ainsi. Selon cette théorie, la signification des nombres figurant dans le calcul des probabilités est la suivante. Dans toute série suffisamment nombreuse de cas semblables, la fréquence relative de l'événement étudié aura la valeur $f : s$ trouvée expérimentalement, ou du moins ne s'en écartera pas sensiblement ; la probabilité mathématique $p(E)$ est une espèce de valeur centrale ou de valeur limite, exactement déterminable ou non, autour de laquelle viennent se grouper les fréquences relatives qu'on trouve expérimentalement. L'interprétation inductive envisage donc l'avenir ; par là, elle va plus loin que l'interprétation empiriste qui se borne au passé.

116. — *Objections soulevées contre l'interprétation inductive.* 1° Il y est fait appel à l'expérience. Or, la base expérimentale fait souvent complètement défaut et néanmoins, les propositions de la théorie des chances restent applicables (v. n° 113).

2° Cette manière d'interpréter postule la nécessité, pour les événements aléatoires, de se dérouler d'une certaine façon. Or, seul un enchaînement de tous les faits suivant des lois inéluctables permet de prévoir l'avenir. La Science postule un déterminisme absolu de tous les événements matériels. En l'admettant, comment expliquer que le côté pile apparaîtra « à peu près aussi souvent » que le côté face, quand on lancera 1 000 fois de suite une bonne pièce de monnaie ? Dans les deux hypothèses faites : 1° que la pièce de monnaie soit bonne (v. n° 38) et 2° qu'on la lance 1 000 fois de suite, il n'y a vraiment rien qui puisse autoriser la conclusion logique que pile apparaîtra nécessairement avec une fréquence relative « à peu près » égale à celle de face.

3° Les théorèmes du calcul des probabilités n'impliquent nullement une pareille nécessité. Au point de vue logique, il reste toujours la possibilité du déroulement contraire ; et la théorie permet même de calculer la probabilité mathématique que dans une série de 1 000 coups à pile ou face, pile apparaisse non pas 500 fois, mais 700 ou 800 ou même 1 000 fois de suite. En effet, cette dernière probabilité, quoiqu'excessivement petite, est assignable ; elle

est $= 1 : 2^{1000} = 0,00\ldots$ (300 zéros)$\ldots$ 107 133. Grâce au théorème de Bernoulli, on peut même aller encore plus loin et calculer combien de séries, chacune de 1 000 coups à pile ou face, devraient être exécutées pour qu'on puisse parier par exemple 9 contre 1 de rencontrer au moins une série de 1 000 coups comportant pile 1 000 fois de suite (v. n° 86). Même la probabilité mathématique la plus infime est essentiellement différente de l'impossibilité, et il y a entre elles un abîme infranchissable (v. n° 30).

Le trait caractéristique des propositions de la théorie des chances est et reste le suivant : Quelque justifiée que soit notre expectative, quelque grande que soit notre présomption de voir un événement aléatoire E se réaliser, cette réalisation n'est jamais une nécessité, la non-arrivée de E demeure toujours possible. Changer l'ordre de grandeur des probabilités mathématiques qui interviennent ne change pas leur nature ; envisager une nombreuse série d'épreuves répétées au lieu d'une observation isolée ne résout pas la difficulté. On peut parfaitement accorder que, très souvent même, les propositions de la théorie des chances expriment un fait expérimental ou ont été trouvées par induction ; mais par cela, leur sens philosophique n'est pas encore élucidé.

Résumons la discussion ci-dessus. Il n'est pas facile de dire d'une manière précise et complète quelles sont les circonstances objectives qui guident nos présomptions et servent de base à de curieuses règles d'expectative, en permettant de statuer des cas équipossibles sans aucun arbitraire. En tout cas, le principe de l'absence de raison justifiante doit être complété ou, mieux encore, remplacé par un autre : celui de la raison suffisante ou même celui de la raison péremptoire. Pour faciliter l'étude des dites circonstances objectives, nous allons poser préalablement quelques définitions importantes.

117. — La connaissance de la réalité est un problème comprenant deux parties essentiellement différentes.

a) Découvrir les lois de la nature. Il s'agit de trouver les principes des actions et des réactions dans l'évolution des phénomènes, les lois régissant les transformations que nous voyons s'opérer continuellement dans l'état de l'univers, tant microscopique que stellaire.

Définition 3. — Pour avoir un nom bref et commode à disposition, j'appellerai ces lois ou principes ou propriétés générales *les circonstances nomologiques* (66) ; en langage mathématique, ce sont les équations différentielles des phénomènes. — *Exemples.* 1) La loi de l'attraction universelle découverte par Isaac Newton : Les phénomènes se passent comme si deux masses matérielles m_1 et m_2 s'attiraient avec une force f directement proportionnelle au produit des masses et inversement proportionnelle au carré de leur distance r ; $f = m_1 . m_2 : r^2$. — 2) La vitesse d'un express est de 90 kilomètres à l'heure. — 3) L'industriel N gagne net 50 000 francs par an. Ces propositions expriment certaines lois qui peuvent d'ailleurs être exactes ou approchées, durables ou éphémères ; rien n'est préjugé à ce sujet par le terme de *nomologique.*

b) Connaître et décrire le point de départ des phénomènes à étudier. En langage mathématique, ce sont les conditions initiales ou conditions aux limites ; en d'autres termes, il s'agit de savoir quel est à un moment précis t_0 l'état exact de l'ensemble dont on veut suivre les transformations successives.

Définition 4. — Par opposition aux circonstances nomologiques ci-dessus définies, j'appellerai cet état initial : *les circonstances ontologiques* (67). — *Exemples.* Reprenons ceux mentionnés sous lettre *a*). 1) Le premier janvier 1900 à minuit, l'état du Système solaire était : puis suivrait une longue liste donnant, pour chacun des corps célestes de notre système planétaire, la masse, la position exacte, la grandeur et la direction de la vitesse dont il est animé. 2) L'express pour Ostende est parti de la gare de Bruxelles à huit heures du matin. 3) La fortune nette de l'industriel N s'élève actuellement à 900 000 francs.

118. — Des définitions précédentes découlent plusieurs corollaires. *a*) On remarquera tout d'abord que ces deux catégories de

(66) C'est-à-dire « indiquant une loi » ; du grec ὁ νόμος = la loi (se retrouve par exemple dans astro-nomie, arith-nomie etc.) et ὁ λόγος = la parole, la doctrine (se retrouve par exemple dans zoo-logie, géologique etc.).

(67) C'est-à-dire « indiquant une manière d'être » ; du grec τὰ ὄντα, participe de εἶναι = être, exister, et ὁ λόγος = la parole, le mot. Voir la note précédente.

circonstances sont indépendantes l'une de l'autre. En particulier, les circonstances ontologiques peuvent changer souvent et très profondément pendant que les nomologiques restent les mêmes. Les conditions initiales admettent une certaine marge, un certain degré de déjour ou d'ébattement, elles peuvent varier infiniment entre certaines limites et présenter toute une gamme de modalités.

Définition 5. — L'ensemble de ces diverses possibilités ontologiques constitue une zone de variabilité, un certain champ de variation, un certain domaine que j'appellerai *zone de comportement* ou *spaciom comportaire*. Il est donc de nature essentiellement ontologique ; les conditions nomologiques restent constantes ou sont supposées invariables.

b) Quiconque veut prévoir l'évolution des phénomènes doit connaître tout aussi bien les circonstances ontologiques que les nomologiques. La connaissance de l'une seulement de ces deux catégories ne permet pas de prédire l'avenir. On s'en rend aisément compte par les exemples sus-mentionnés. Ainsi, 1° un astronome aurait beau connaître les lois de Képler, de Newton et tous les théorèmes de la mécanique (circonstances nomologiques), il serait incapable de prédire une éclipse s'il ne connaissait pas en même temps, pour une époque déterminée d'ailleurs arbitrairement choisie, les positions relatives du Soleil, de la Terre et de la Lune (circonstances ontologiques). De même 2°, on aurait beau connaître exactement la loi de variation de la vitesse du train en fonction du temps (circonstance nomologique), on ne pourrait pas donner une réponse exacte à la question : « Dans combien d'heures l'express arrivera-t-il à Ostende? » si l'on ne savait pas du tout où il se trouve au moment dont il s'agit (circonstance ontologique). De même encore 3°, bien que connaissant l'augmentation annuelle nette de la fortune, il serait impossible de répondre exactement à la question : « Dans combien d'années l'industriel *N* sera-t-il devenu millionnaire? » si l'on ignorait totalement l'état actuel de sa fortune.

c) On voit que les conditions nomologiques présentent toujours un certain caractère de généralité, en ce sens qu'elles sont valables pour toute une catégorie de cas ou d'exemplaires, pour tous les

éléments d'un certain ensemble ; elles sont abstraites. Les circonstances ontologiques, au contraire, sont toujours individuelles, valables seulement pour un cas déterminé ou pour 1 élément ; elles sont concrètes et non réductibles à des nécessités générales.

119. — La distinction entre ce qui est ontologique et ce qui est nomologique se retrouve avec beaucoup de force dans la notion souvent employée de « *possibilité objective* ou physique ». Elle est essentiellement différente de cette possibilité, toute subjective, qui résulte d'une ignorance totale et de l'absence complète de connaissances touchant l'événement étudié. Par exemple, lorsqu'on dit qu'il est « objectivement possible » d'amener dix fois de suite le point 6 avec un dé à jouer, on entend par là que cet événement composé n'est pas contraire aux conditions nomologiques, aux lois de la nature. C'est la même chose que la « possibilité métaphysique » chez Kant. On veut dire que les circonstances ontologiques : la façon d'agiter le cornet, d'en faire sortir le dé etc. etc., circonstances variables à l'infini, n'impliquent rien qui exclurait la réalisation du dit événement. — Au lieu du dé à jouer, mettez dans le cornet un bouchon de liège. Quelle que soit la manière de le lancer dans l'eau, il n'ira jamais au fond du baquet ou du lac. En disant que cela est « objectivement impossible », on veut dire que c'est contraire aux lois de la nature, donc aux conditions nomologiques.

a) En vous montrant un corps que vous ne connaissez pas du tout, on vous demande s'il ira au fond de l'eau. Si vous répondez : « C'est possible », parceque vous ignorez totalement son poids spécifique, cette possibilité-là, n'étant pas basée sur des connaissances nomologiques, sera purement subjective. Elle ne pourrait donner lieu qu'à des probabilités *anumériques*.

On voit que ces diverses notions : circonstances ontologiques et nomologiques d'une part, possibilité objective et subjective d'autre part, peuvent s'éclairer l'une l'autre et jeter de la lumière sur les notions de probabilité numérique et de probabilité *anumérique*. En effet, une condition nécessaire pour qu'il puisse y avoir probabilité numérique est l'existence de circonstances nomologiques connues ou du moins tacitement admises. Le principe de la raison suffisante l'exige. L'ignorance de celui qui conjecture, si ignorance

il y a, doit porter exclusivement sur les circonstances ontologiques (voir les numéros 120 *b* et 154).

b) D'une façon générale, la critique philosophique du calcul des probabilités s'est attaquée au vocabulaire, en particulier à ce terme de « cas également possibles » ; doit-on dire plutôt « également probables » ou « également vraisemblables » ? Y a-t-il lieu de faire une différence entre ces synonymes ? Et laquelle ? La seule objection sérieuse me paraît être la suivante. Une chose est ou bien possible ou impossible ; une troisième alternative n'existe pas. Quiconque répondra par la négative, à la question si l'on peut voir quelque chose de quantitatif dans la possibilité, devra rejeter le terme « également possibles ».

Pour éviter ces critiques, beaucoup d'auteurs ont employé d'autres termes ; par exemple Mr de Kries [64] parle de « suppositions également bien fondées » (« gleichberechtigte Annahmen »). Je dirai de préférence *équiprobables* ou également vraisemblables. Mais le terme « également possibles » ou *équipossibles* a acquis droit de cité dans le langage, et l'on doit dès lors admettre une possibilité susceptible de gradation. Sur elle se concentre l'intérêt philosophique.

VI. — *Le principe des zones de comportement.*

120. — Dans quelles conditions ou circonstances objectives peut-on statuer, d'une manière excluant tout arbitraire, des cas équiprobables ou également possibles ? Voilà ce qu'il s'agit de déterminer. Cherchant à généraliser une idée suggérée par l'exemple 2 du numéro 109, on présumera qu'il y a toujours, à la base de cas équipossibles, une certaine étendue globale qui se décompose en parties mesurables ou du moins comparables entre elles. Je désignerai dans la suite ce *spaciom global* par *Sp* (v. n° 118, *a*).

a) Dans l'exemple du bolide tombant sur la Terre, la probabilité pour un pays d'être touché étant posée proportionnelle à sa surface, si l'on divise la surface du globe terrestre en n parties égales, on aura n cas équipossibles à condition qu'il n'y ait aucune raison d'ordre logique pour considérer l'un comme plus probable qu'un

autre. La présomption ou l'expectative doivent se baser exclusivement sur des grandeurs objectives, en particulier sur des étendues bien déterminées. Pour exprimer brièvement ce fait, adoptons les définitions suivantes.

Définition 6 (v. n° 118 *a*). — Des zones de comportement ou spacioms comportaires sont dits *indifférents*, s'il n'existe aucune raison d'ordre logique pour préférer l'un à l'autre ou pour considérer l'un plutôt que l'autre.

Définition 7. — Une expectative ou présomption ou conjecture est dite *libre*, si elle est fondée non pas sur des considérations d'ordre logique ou psychologique, mais uniquement sur des zones comportaires indifférentes dans le sens de la définition 6.

b) On voit facilement que la connexion logique ne présente pas toujours le caractère de l'expectative libre dans le sens de la définition 7. Par exemple dans le raisonnement par analogie : on se demande jusqu'à quel point il y a similitude du cas actuel avec ceux antérieurement observés et l'on pèse l'importance, physiologique ou autre, des ressemblances et des différences ; on fait donc entrer en jeu autre chose que des grandeurs physiquement mesurables ; ce n'est donc plus une conjecture « libre » ; elle ne saurait donner naissance à des probabilités mathématiques. Ainsi, la probabilité qu'un cas de fièvre typhoïde ait une issue heureuse est une probabilité *anumérique*.

c) On verrait qu'il en est de même dans le raisonnement par induction (v. n° 114). Pour évaluer la probabilité de l'exactitude d'une loi générale de la nature, par exemple de la loi de la constance de l'énergie, on considère, outre le nombre et l'importance des vérifications expérimentales, encore sa fécondité en nouvelles prévisions, les critiques qu'elle a suscitées, les objections qu'elle pourrait soulever etc., en un mot : on fait entrer en jeu des considérations d'ordre logique. La présomption n'est donc pas libre dans le sens de la définition 7 et la probabilité qu'on en déduirait ne serait pas une probabilité mathématique.

121. — Pour l'existence de probabilités mathématiques, les conditions du numéro précédent sont nécessaires, mais non suffi-

santes ; l'exemple suivant le prouvera. Je vois deux cartes de jeu placées l'une à côté de l'autre. Je n'ai aucune idée comment ni pourquoi elles se trouvent là. J'en retourne l'une et je constate qu'elle est noire. Quelle est la probabilité α pour que l'autre soit rouge ?

A première vue, on dira que la seconde carte peut tout aussi bien être rouge que noire, donc $\alpha = \frac{1}{2}$. Mais Poisson, dans ses *Recherches sur la probabilité des jugemens...* (3), p. 96, en se basant sur le théorème de Bayes, arrive à la conclusion que la dite probabilité α est $= 1/3$.

On peut aussi dire qu'il y a à priori quatre cas également vraisemblables : *a*) les deux cartes noires ; *b*) la première noire et l'autre rouge ; *c*) la première rouge et l'autre noire ; *d*) les deux cartes rouges. La possibilité *d*) se trouve écartée, puisque je sais par expérience que l'une des deux cartes est noire ; restent donc trois cas, puisque, ne sachant pas dans quel ordre de succession les cartes ont été tirées et placées là, j'ignore quelle est la première, quelle est la seconde. De ces trois cas équipossibles, les deux derniers *b*) et *c*), sont favorables à l'événement considéré. La probabilité cherchée est donc $\alpha = 2/3$.

Ce résultat serait exact *si je savais* que les cartes ont été tirées au hasard, de deux jeux égaux contenant chacun autant de noires que de rouges. Mais comme j'ignore s'il en est ainsi, je dois en supputer la probabilité p_0, arbitraire d'ailleurs. J'aurai alors un événement composé, et les probabilités des quatre cas *a*), *b*), *c*) et *d*) seront fonction de ce p_0 ; de même la probabilité cherchée α.

Je puis aussi supposer que les deux cartes ont été extraites, au hasard, d'un même jeu et calculer, dans cette hypothèse, les probabilités des quatre cas en question. Elles seront différentes suivant que je supposerai les cartes extraites d'un jeu de Whist à 52 cartes ou d'un jeu de Piquet à 32 cartes etc. Je trouverai pour α de nouvelles valeurs correspondant à ces hypothèses. Comme j'ignore complètement laquelle est réalisée, je devrai évaluer, arbitrairement, les probabilités p_1, p_2, ..., que les deux cartes ont été extraites, au hasard, d'un jeu de 52 ou de 32 etc , puis appliquer le théorème de multiplication (v. n° 34).

Ce n'est là que le commencement d'une série à peu près illimitée de possibilités. D'ailleurs, qui me dit que les cartes ont été tirées au hasard et qu'on ne les a pas choisies intentionnellement ? On voit bien que si l'on voulait aller au fond de la question, on s'engagerait dans un labyrinthe de possibilités diverses, chacune affectée d'une probabilité plus ou moins arbitraire dont il serait bien difficile de tenir un compte exact. Conclusion : dans les conditions admises, la probabilité α est tout-à-fait indéterminée.

122. — L'exemple précédent suffit à prouver que la probabilité mathématique n'est pas nécessairement assignable quand elle résulte d'une conjecture libre dans le sens de la définition 7. Il faut que la zone de comportement envisagée soit encore *primaire* ; autrement dit, la probabilité d'un état présent ou futur doit être évaluée suivant la probabilité des états antérieurs qui le conditionnent. Ceux-ci étant ou pouvant être à leur tour conditionnés par d'autres, antérieurs à eux, on devra remonter à ceux-là et ainsi de fil en aiguille, jusqu'à un état qui ne soit plus réductible à d'autres, qui soit primordial.

Définition 8. — C'est cet état primordial que nous appellerons *primaire*. On a un état primaire, si l'on aboutit à des grandeurs constantes, par exemple au poids spécifique d'une substance.

Pour qu'il soit possible d'attribuer un sens à une probabilité numérique, l'état de nos connaissances doit donc permettre de statuer des zones de comportement qui soient : comparables entre elles, indifférentes et primaires. De plus, on doit pouvoir épuiser le domaine global Sp, c'est-à-dire l'ensemble de toutes les possibilités entrant en question, par un nombre fini de tels spacioms comportaires ; chacun d'eux doit correspondre à l'un des résultats finaux possibles (v. n° 118, définition 5). On aboutit ainsi à la

Définition 9. — Des cas sont *équipossibles* ou *également vraisemblables*, s'il leur correspond, dans le sens des définitions 6 et 8, des domaines de comportement indifférents, primaires et égaux.

Définition 10. — Appelons le principe auquel les considérations ci-dessus nous ont conduit *le principe des zones de comportement*

ou *principe des spacioms comportaires.* Il est très général et naturellement très important pour les fondements logiques et philosophiques de la théorie des chances, quand on se place sur le terrain de la définition classique.

Corollaire 1). — Si un cas ou un événement est tel que le champ comportaire correspondant comprenne la presque totalité du domaine global *Sp* qui entre en question (v. n° 120), cet événement a pour lui une très grande vraisemblance, sa probabilité mathématique est très voisine de 1.

Corollaire 2). — Il y a deux espèces de primordialité permettant d'assigner un sens à des probabilités numériques. Les voici.

a) Lorsque, en remontant de fil en aiguille la série des états successifs correspondant aux époques t_1, t_2, t_3, ..., on finit par arriver soit à une grandeur constante (par exemple au poids spécifique d'une substance ou à un autre coefficient physique), soit à une origine des temps, mettons à l'époque t_n, à partir de laquelle on puisse affirmer que les facteurs considérés sont entrés en jeu. La probabilité mathématique sera $= u_n : A_n$ si, pour la dite époque t_n, A_n désigne l'étendue du spaciom global *Sp* (v. n° 120) et u_n la mesure du domaine comportaire de l'événement étudié. C'est *la primordialité de première espèce.*

b) Lorsque, une origine des temps n'étant pas assignable, le rapport $u_\lambda : A_\lambda$ conserve toujours la même valeur,

$$\frac{u_1}{A_1} = \frac{u_2}{A_2} = \frac{u_3}{A_3} = \ldots = \frac{u_\lambda}{A_\lambda} = \ldots = \text{constante} \equiv p.$$

Dans ce cas encore, on pourra parler de la probabilité mathématique p, car on sera fondé à appeler « primaire » l'état actuel, puisque les cas actuellement équipossibles le restent toujours, même si l'on remonte aux circonstances dont ils sont les conséquences nécessaires en vertu des lois de la nature. C'est *la primordialité de seconde espèce.*

Il ressort de ces considérations que la primordialité des domaines comportaires est une notion philosophique très importante.

La théorie des cas équipossibles n'est solidement fondée que si elle se base sur des domaines de comportement comparables entre

eux, indifférents dans le sens de la définition 6 et primaires. Le principe de l'absence de raison justifiante est insuffisant comme fondement philosophique d'une discipline scientifique.

Le jeu de la roulette rectiligne

123. — Le jeu suivant, étudié par Mr Jean de Kries ([68]) et que j'appellerai *la roulette rectiligne*, est propre à éclaircir la théorie des cas équipossibles. Qu'on imagine une très longue rainure ou rigole horizontale dont le fond, bien lisse et bien plan, soit strié perpendiculairement au sens de la longueur, c'est-à-dire divisé en bandes transversales parallèles, alternativement rouges et noires. Pour fixer les idées, supposons chaque bande large de 1 millimètre. A l'extrémité de cette rainure, on place une bille ou une paume qu'une impulsion fait rouler sur les stries transversales et qui parcourt un certain chemin, mettons l centimètres. Je dirai pour abréger qu'on a « amené strie noire », si la bille finit par s'arrêter sur une strie noire.

Question. Quelles sont : la probabilité β d'amener strie noire et la probabilité γ d'amener strie rouge? Il résulte manifestement de nos hypothèses que la probabilité d'amener strie rouge est la même que celle d'amener strie noire, $\beta = \gamma$. Et comme ces deux résultats sont exclusifs l'un de l'autre et que l'un d'eux se réalise nécessairement, $\beta + \gamma = 1$ (v. n° 23). Dès lors, on a la

Réponse : $\beta = \gamma = 0{,}5$ sont les probabilités cherchées.

Examinons de plus près pourquoi ces deux résultats, strie noire et strie rouge, sont ou paraissent également possibles.

a) L'équivraisemblance ne repose pas, contrairement à ce que l'on pourrait penser au premier abord, sur l'idée que la probabilité d'amener n'importe quelle strie soit exactement la même; on peut fort bien admettre que les probabilités d'amener la strie numéro 500 et la strie numéro 10 000 sont très différentes. Il faut aussi renon-

(68) Voir l'ouvrage cité dans la note (64), au bas de la page 160, notamment le chapitre III, § 2 : *das Stoss-Spiel* = « le jeu du choc » ou « le jeu de la [boule] poussée ».

cer à essayer de comparer numériquement les probabilités de deux cas aussi éloignés l'un de l'autre, dont le premier résulte d'une impulsion très modérée et le second d'une impulsion très forte communiquée à la bille.

b) Il suffit d'admettre que les probabilités d'amener deux stries contiguës, la $n^{\text{ième}}$ qui est rouge et la $(n+1)^{\text{ième}}$ qui est noire, sont égales ou du moins ne diffèrent que très peu l'une de l'autre, et que la probabilité pour la bille lancée de s'arrêter sur une bande de largeur b est proportionnelle à b. Additionnons les largeurs de toutes les stries noires d'une part, de toutes les stries rouges d'autre part, parcourues par la bille ; nous obtiendrons deux sommes égale chacune à $l : 2$, donc égales entre elles. Dès lors, en appliquant le théorème des probabilités totales (v. n° 32), on se rend compte que la somme des probabilités d'amener strie rouge et la somme des probabilités d'amener strie noire sont égales entre elles ou ne peuvent différer que d'une quantité négligeable.

124. — Pour être un jeu de hasard dans le sens classique du terme, la roulette rectiligne doit remplir encore d'autres conditions : *une très légère variation dans les conditions initiales*, dans l'intensité ou dans la direction ou dans le point d'application de la force qui fait rouler la bille, *doit suffire pour que le résultat final soit autre*, strie rouge au lieu de strie noire.

a) Il est essentiel tout d'abord que les stries rouges et noires qui alternent régulièrement soient non seulement très nombreuses et de même largeur, mais que cette largeur soit petite par rapport au chemin l que parcourt la bille avant de s'arrêter. Si au lieu d'un millimètre, comme nous l'avons supposé ci-dessus, chaque strie avait une largeur de dix mètres par exemple, le nombre des bandes qui entreraient en question pour un coup serait très restreint, peut-être deux ou trois seulement, et l'on ne pourrait plus poser à peu près égales entre elles les probabilités que la bille s'arrêtera sur une strie rouge ou sur la strie noire contiguë. Pour évaluer la probabilité d'amener strie noire, il faudrait faire entrer en ligne de compte l'intensité de l'impulsion communiquée à la bille, circonstance ontologique qui introduirait des probabilités subjectives et anumériques (v. n° 117, *b* et n° 119, *a*).

b) Il faut encore que la bille ne soit pas très lourde ni très adhérente en comparaison de la force qu'on lui applique pour la faire rouler ; si par exemple une force de 15 365 dynes parallèle à l'axe de la rainure a amené ou amène toujours strie noire, il faut que 15 364 ou 15 366 dynes parallèles à l'axe de la rainure amènent strie rouge et que la plus petite modification dans la direction ou dans le point d'application change également le résultat final. C'est ce qu'on exprime en disant : il faut qu'il y ait hasard.

Définition 11. — Il y a *hasard subjectif*, lorsqu'une petite cause produit de grands effets ; la cause insignifiante n'est pas perçue, mais l'effet nous frappe et nous paraît fortuit.

c) Enfin, *il faut que les circonstances ontologiques soient indifférentes pour l'expectative*. Ce ne serait pas le cas si, par exemple, au commencement et à la fin de la rainure les stries noires étaient toujours trois fois plus larges que les rouges, tandisque dans la partie centrale il y aurait égalité de largeur de toutes les stries. Il ne serait alors plus indifférent, pour notre expectative, de supposer une impulsion faible ou moyenne ou forte. Les diverses régions du domaine global *Sp* cesseraient d'être équivalentes pour l'estimation de la probabilité, et certaines conditions initiales devraient entrer en ligne de compte. L'indifférence des circonstances ontologiques repose donc sur les faits suivants : 1° Chacun des deux domaines comportaires possibles, celui qui amène strie noire et l'autre qui amène strie rouge, est dispersé en un très grand nombre de spacioms partiels ; ceux-ci, grâce à leur alternance régulière et systématique, constituent un mélange homogène dans toutes les parties du domaine global *Sp* entrant en question. 2° Les rapports numériques des domaines partiels contigus sont constants et les mêmes partout.

125. — Il résulte immédiatement des considérations précédentes que les deux zones de comportement possibles dans ce jeu sont égales et indifférentes dans le sens de la définition 6. Je dis qu'elles sont aussi primaires. En effet, c'est le cas par définition si l'on s'arrête à l'instant précis t où l'impulsion est communiquée à la bille ; on a la primordialité de première espèce (v. n° 122, corollaire 2). Et si l'on remonte le cours du temps, qu'on envisage les

causes du roulement de la bille, donc les mouvements de la main lui communiquant l'impulsion ou les contractions musculaires ou l'action des nerfs aboutissant à ces mouvements ou les causes très complexes et très nombreuses qui ont amené l'opérateur à lancer la bille, et ainsi de suite, on constatera ceci : à tout ensemble M_n de causes, de raisons et de circonstances qui finalement ont amené ou amèneraient strie noire, on peut faire correspondre un second ensemble, M_r, qui diffère très peu du premier, qui est contigu à M_n, mais qui finalement aurait amené ou amènerait strie rouge ; et la somme de tous les M_n est égale à celle de tous les M_r ou n'en diffère que très peu ; et ce fait se maintient, quelle que soit l'époque à laquelle on s'arrête en remontant le cours du temps. On a donc la primordialité de seconde espèce.

Remarque. — Si l'on ne savait absolument rien de la largeur des stries, il n'y aurait dans l'état de nos connaissances aucune raison d'admettre que la bille s'arrêtera sur une bande noire plutôt que sur une rouge ; on pourrait poser $\beta = \gamma = \frac{1}{2}$ en invoquant le principe de l'absence de raison justifiante. Ce serait fonder le calcul des probabilités sur une base trop peu solide. Vouloir tirer des conclusions scientifiques de prémisses aussi vagues serait vouloir créer de la science avec de l'ignorance ; c'est impossible. Il faut substituer à ce principe celui de la raison péremptoire (voir les numéros 19, 111 et 143).

126. — Il est facile de généraliser le jeu de la roulette rectiligne. *a*) Si l'on savait que tout le long de la rainure, les stries rouges sont toujours 2 fois ou 3 fois ou généralement r fois plus larges que les stries noires, tout en alternant régulièrement avec elles, on serait fondé à dire, voire même obligé de conclure, qu'il y a 2 fois ou 3 fois ou généralement r fois plus de chances pour que la bille finisse par s'arrêter sur une strie rouge plutôt que sur une noire (principe de la raison péremptoire). On en déduirait que la probabilité mathématique β d'amener strie rouge est 2 fois ou 3 fois ou généralement r fois plus grande que celle d'amener strie noire, γ ;

et puisque $\beta + \gamma = 1$, on aurait les valeurs $\beta = 2/3$ et $\gamma = 1/3$ ou $\beta = 3/4$ et $\gamma = 1/4$ ou généralement

$$\beta = \frac{r}{r+1} \text{ et } \gamma = \frac{1}{r+1}.$$

b) *Problème.* — Si les stries de notre rainure étaient coloriées en plus de deux couleurs, qu'elles fussent par exemple alternativement rouges, blanches, bleues et noires, se succédant toujours dans le même ordre et toutes de même largeur, quelles seraient : la probabilité mathématique α d'amener strie rouge? la probabilité β d'amener strie blanche? la probabilité γ d'amener strie bleue? et la probabilité δ d'amener strie noire? *Solution.* — On sera fondé à regarder comme équipossible l'arrêt de la bille sur l'une quelconque des quatre couleurs et à déduire de là $\alpha = \beta = \gamma = \delta = 1/4$, uniquement si l'on *sait* que les stries sont exactement de même largeur, et cela dans toutes les parties de la rainure (v. n° 124, *c*), et la remarque finale du numéro 125). — Si par contre on savait que, les stries rouges et les stries noires par exemple étant toujours de même largeur, les blanches sont deux fois et les bleues cinq fois plus larges que les rouges, et cela dans toutes les parties de la rigole, les probabilités mathématiques d'amener ces couleurs seraient :

$$\alpha = 1/9,\ \beta = 2/9,\ \gamma = 5/9 \text{ et } \delta = 1/9.$$

L'extension au cas de n couleurs ne présente pas de difficulté aux points de vue philosophique, logique et mathématique, pourvu que les rapports des largeurs soient constants, connus et les mêmes dans toute l'étendue du domaine global *Sp* (voir le numéro 123 et le numéro 124, *c*).

On pourrait répéter ici la *Remarque* du numéro précédent. Si l'on ne savait rien de la largeur respective des stries, il serait très arbitraire de poser à priori $\alpha = \beta = \gamma = \delta = 1/4$ en se basant sur le principe de l'absence de raison justifiante. Les conclusions qu'on voudrait tirer de cette égalité seraient illusoires.

127. — Dans le jeu de la roulette rectiligne, l'existence des probabilités mathématiques se ramène au principe des zones de comportement. En examinant à ce point de vue des cas plus compliqués, on arrive à la conclusion qu'il en est de même lorsqu'il

s'agit du jeu du dé ou de cartes à jouer ou du tirage d'un billet de loterie ou du schéma de l'urne. On prouve d'abord que le domaine global *Sp* qui entre en question se subdivise en *n* zones de comportement, dont chacune correspond à l'un des résultats finaux possibles. Dans le cas du dé ordinaire, $n = 6$; dans la loterie génoise, $n = 90$; dans le schéma de l'urne, $n =$ au nombre des couleurs ; etc.

On démontre ensuite que ces *n* spacioms comportaires sont indifférents dans le sens de la définition 6, comparables entre eux et même égaux, parceque chacun se décompose en un nombre énorme de domaines partiels entre lesquels on peut établir une correspondance univoque et réciproque (v. n° 125). On démontre enfin qu'ils sont primaires dans le sens de la définition 8 du numéro 122, pourvu que certaines conditions soient remplies. — Dans le cas de la loterie génoise, la main plongée dans l'urne doit se trouver chaque fois devant un arrangement fortuit de numéros ; sinon, l'arrangement qui existait avant le tirage, donc une circonstance ontologique, devrait être pris en considération ; on aurait peut-être des raisons d'ordre logique de considérer l'un des spacioms comportaires comme plus accessible que les autres ; il n'y aurait donc plus entre eux la relation d'indifférence logique nécessaire à l'existence de probabilités mathématiques. — Dans le cas du jeu de cartes, il est facile de démontrer que certains ordres de succession voisins de l'ordre initial sont beaucoup plus probables que d'autres, si le battage n'a pas été fait assez à fond ; de même dans la loterie, les numéros au fond de l'urne ont moins de chances de sortir que les autres, s'ils n'étaient brassés mécaniquement et bien mélangés à nouveau après chaque tirage. L'égale possibilité qui est due, dans le jeu de la roulette rectiligne, à l'égale largeur des stries, dans le cas du dé à la régularité géométrique et physique parfaite du dé utilisé, provient ici d'un brassage consciencieux et prolongé ; aussi les règles du jeu prescrivent-elles ces manipulations préliminaires.

Si ces conditions sont toutes remplies, il en découle la primordialité des zones de comportement, car les circonstances ontologiques sont alors sans influence sur les rapports numériques de ces zones, même si l'on remonte le cours du temps (v. n° 125). De plus, il suffit d'une variation imperceptible des conditions initiales pour

modifier le résultat final ; l'expectative basée sur ces rapports numériques étant alors libre dans le sens de la définition 7, les probabilités numériques qu'on en déduit ont quelque valeur objective.

128. — *a*) Le principe général des zones de comportement répand beaucoup de clarté sur le théorème de Jacques Bernoulli et sur la nature de la loi des grands nombres. Rappelons que cette loi revient à la proposition suivante. Si l'on sait que le complexe des circonstances qui produisent un événement aléatoire E reste constant ou ne varie que très lentement, et si la probabilité mathématique de E est $= 1/n$ par exemple, on peut être certain que dans une très longue série d'épreuves ou d'observations, le nombre de fréquence de E sera exactement ou « à peu près » la $n^{ième}$ partie du nombre total des épreuves. A cette constatation expérimentale correspond la première partie du théorème de Jacques Bernoulli. Le principe des zones de comportement explique le pourquoi de ce fait. On peut en effet démontrer rigoureusement que l'expectative qui se rapporte à un grand nombre de répétitions du même événement repose sur ce même principe.

Prenons le cas le plus simple et considérons n éléments : a_1, a_2, a_3, .., a_n. Imaginons qu'on en ait formé tous les arrangements possibles r à r avec répétition (v. n° 45), où r représente un grand nombre que, pour simplifier, nous supposerons être multiple de n, par exemple $r = r'.n$; on sait que le nombre de ces arrangements est $n^r = n^{n\,r'} = (n^n)^{r'}$. Soit a_i l'un des n éléments donnés, arbitrairement choisi mais fixe. Je dis que parmi ces arrangements r à r, le nombre de ceux qui contiennent l'élément a_i exactement ou à peu près r' fois, mettons de $(r' - \gamma)$ à $(r' + \gamma)$ fois, forment l'immense majorité. Et le rapport du nombre des arrangements jouissant de cette propriété, au nombre total des arrangements possibles r à r tend vers 1 quand r augmente, d'autant plus rapidement que γ est plus grand.

b) Envisageons un exemple concret, le jeu de croix ou pile (v. n° 38) ou celui de la roulette rectiligne (v. n° 123). Supposons qu'on fasse une série de r coups et qu'on aligne des jetons rouges et des jetons noirs d'après la règle suivante. A chaque coup ame-

nant pile ou strie rouge, on place un jeton rouge ; à chaque coup amenant croix ou strie noire, on place un jeton noir. On forme de cette façon un arrangement r à r avec répétitions de 2 éléments ; on peut en former 2^r différents et chacun d'eux correspond à une série de r coups (v. n° 45). Prenons d'abord $r = 100$. En désignant par p l'arrivée de pile ou de strie rouge et par c celle de croix ou de strie noire, on constate ceci : Parmi les $2^{100} \sim 126\,765 . 10^{25}$ séries possibles de 100 coups à p ou c, celles qui contiennent exactement ou à peu près autant de fois p que c forment l'écrasante majorité. En effet, le nombre des combinaisons possibles 100 à 100 contenant 50 p et 50 c est $\binom{100}{50} = \frac{100!}{50!\,50!}$, et le nombre des combinaisons 100 à 100 contenant par exemple 90 p et 10 c est $\frac{100!}{90!\,10!}$. Le rapport de ces deux nombres est $= \frac{90!\,10!}{50!\,50!}$. En vertu de la formule de Stirling (v. n° 47), ce rapport devient approximativement $3^{181} : 5^{101} = 5\,794\cdot\cdot 000\,000\cdot 000\,000$. Pour $r = 100$, le premier nombre est donc presque six millions de milliards de fois plus grand que le second. L'ordre de grandeur n'est pas modifié sensiblement si, au lieu de 50 p + 50 c, on prend les séries de 49 p + 51 c ou celles de 48 p + 52 c ou de 47 p + 53 c ou présentant une fluctuation du même ordre. On voit par là que sur les 2^{100} arrangements possibles, le nombre de ceux qui contiennent exactement ou à peu près autant de jetons d'une couleur que de l'autre est des millions de milliards de fois plus grand que le nombre des autres. En appliquant alors le théorème d'addition (v. n° 32), on comprend le pourquoi de la loi des grands nombres.

Au lieu de $r = 100$, prenons $r = 1\,000$; le résultat précédent sera très fortement accentué. Le nombre des séries différentes possibles de 1 000 coups à croix ou pile est $2^{1000} \sim 10\,713 . 10^{297}$, donc un nombre de 302 chiffres. Le nombre des combinaisons possibles 1 000 à 1 000 contenant 500 p + 500 c est $\frac{1\,000!}{500!\,500!}$, et de celles contenant 900 p + 100 c, le nombre est $\frac{1\,000!}{900!\,100!}$. Le rapport de ces deux nombres devient, par la formule de Stirling, approximativement $3^{1801} : 5^{1001} \sim 42\,315 . 10^{155}$, soit un nombre de

160 chiffres. Et l'ordre de grandeur de ce rapport n'est pas modifié sensiblement si, au lieu de $500\,p + 500\,c$, on prend les séries de $501\,p + 499\,c$ ou de $502\,p + 498\,c$ ou de $503\,p + 497\,c$, etc. Ce sont là des résultats immédiats d'un dénombrement rigoureux.

129. — Les exemples du numéro précédent prouvent déjà que, comparé au nombre des séries contenant exactement ou à peu près autant de fois p que c, le nombre des autres, c'est-à-dire de celles qui présentent un écart un peu considérable, est tout-à-fait infime et disparaît d'autant plus complètement que r est plus grand. On peut en tirer plusieurs conséquences dont nous mentionnons ici les deux suivantes. — *a*) Une fluctuation un peu considérable a une probabilité mathématique excessivement faible et tout-à-fait négligeable, dès que le nombre d'épreuves est grand. — *b*) Le principe philosophique qui sert de base théorique et de point de départ à nos conjectures est toujours le même, c'est le principe des zones de comportement. Qu'il s'agisse non plus d'un fait isolé, mais d'un événement aléatoire E poursuivi à travers de longues séries d'épreuves répétées, je dis que c'est encore le principe des spacioms comportaires qui guide notre expectative, quand nous nous attendons à voir E se répéter avec une fréquence relative conforme à ce que prédit le théorème de Jacques Bernoulli. En effet, si l'on prend comme événement considéré « amener pile exactement ou à peu près $r/2$ fois », le spaciom comportaire correspondant, en vertu du théorème d'addition du calcul des probabilités (v. n° 32), est mesuré directement par le nombre des séries à r coups dans lesquelles cette fréquence se trouve réalisée ; et l'on vient de voir que ce nombre égale presque 2^r, nombre total des séries possibles.

Supposons qu'au lieu de $n = 2$ modalités comme tout à l'heure (rouge ou noir, croix ou pile), il y en ait un plus grand nombre, par exemple $n = 6$ comme dans le cas du dé ordinaire. Envisageons l'ensemble des 6^r séries différentes possibles à r coups. Celles d'entre elles qui contiennent exactement ou à peu près le même nombre de fois chacun des six points, donc chacun d'eux exactement ou à peu près $r : 6$ fois, forment l'immense majorité, d'autant plus écrasante que r est plus grand. On le démontrerait par un raison-

nement analogue au précédent, fondé sur l'analyse combinatoire.

Même résultat dans le cas général où le nombre n des modalités ou des résultats finaux possibles est quelconque. Dans l'ensemble des n^r séries possibles à r coups, la proportion de celles qui renferment exactement ou à peu près $r : n$ fois chacun des résultats finaux possibles est tellement forte qu'il ne reste pour les autres qu'une marge infime, d'autant plus évanouissante que le nombre d'épreuves est plus grand.

En résumé, les déductions précédentes prouvent deux choses :

1° Qu'il s'agisse d'un événement aléatoire isolé ou du nombre de fréquence d'un même événement soumis à des épreuves répétées dans des circonstances identiques, les conjectures relatives à son arrivée sont toujours de même nature ; dans les deux cas, elles reposent sur le même principe, celui des zones de comportement. Mais dans les deux cas, on ne connaît non plus rien qui exclurait logiquement le résultat contraire ; tout ce qu'on peut dire de lui, c'est qu'il est excessivement peu vraisemblable.

2° Les conjectures et expectatives sont d'autant plus solidement fondées que le nombre r des épreuves ou des observations est plus grand. Cela tient au fait que les séries favorables, c'est-à-dire contenant exactement ou à peu près le même nombre de fois chacun des n résultats finaux possibles, forment, dans le total de toutes les séries à priori possibles, une majorité d'autant plus énorme que r est plus grand.

C'est l'égalité des zones de comportement qui entraîne l'égalité des fréquences relatives, et cela dans les domaines les plus divers. Il semble y avoir là une loi naturelle d'une très grande universalité.

CHAPITRE VI

LA PROBABILITÉ MATHÉMATIQUE BASÉE SUR LA THÉORIE DES ENSEMBLES

I. — *Considérations préliminaires.*

130. — Par des voies abstraites et originales différentes de celles exposées dans le chapitre V, Mr A. Meinong arrive à des résultats qui concordent en quelques points essentiels avec les conclusions développées dans les pages précédentes. Comme elles éclairent le problème sous un autre jour, nous allons reprendre quelques-unes des notions introduites par lui dans son ouvrage *Sur la possibilité et la probabilité* (69). La plus importante est la distinction entre « chose complète » et « chose incomplète », le mot de chose étant pris dans le sens le plus large possible : tout ce qui est accessible à la pensée humaine.

Définition 1. — Une *chose complète* est caractérisée par le fait qu'on peut toujours répondre par oui ou par non, à toute question si telle et telle propriété ou attribut ou relation appartient à la chose ; et si l'on ne peut pas toujours donner cette réponse déterminée, si l'on doit répondre parfois « Cela dépend », la chose est dite *incomplète.*

Exemples. — 1) Tout objet concret, le fourneau que je vois devant moi, la table sur laquelle j'écris etc., de même aussi toute sensation perçue, tout plaisir vécu, sont des choses complètes, car on peut répondre par oui ou par non à toute question de propriété

(69) A. Meinong, *Ueber Möglichkeit und Wahrscheinlichkeit.* Beiträge zur Gegenstandstheorie und Erkenntnistheorie, Leipzig, 1915. Voir aussi E. Czuber, *Die philosophischen Grundlagen der Wahrscheinlichkeitsrechnung,* Leipzig und Berlin, 1923.

ou de relation s'y rapportant : est-il rouge ? est-il rond ? est-il métallique ? durable ? de tel et tel poids ? à telle distance de cet autre objet ? etc.

2) Toute notion est une *chose incomplète* dans le sens de la définition 1. Par exemple la notion de triangle ; car à la question « Est-il rectangle ? » « Est-il isocèle ? » je ne puis pas répondre par oui ou par non. De même la notion abstraite d'arbre; l'arbre en soi n'est ni petit ni grand, ni sec ni florissant ; c'est une chose incomplète; mais l'arbre particulier que je vois devant moi ou celui auquel je pense en ce moment est nécessairement grand ou petit, sec ou florissant; c'est une chose complète. De même le triangle particulier que je vois tracé sur le papier ou celui auquel je pense maintenant est nécessairement rectangle ou non, équilatéral ou isocèle ou scalène etc. ; c'est une chose complète.

L'intelligence humaine est incapable de saisir des choses complètes, à cause de la multitude infinie des déterminations qui font précisément qu'elles sont complètes. Avoir présentes à l'esprit toutes les propriétés d'une chose complète et toutes les relations de distance, de simultanéité, de succession, de similitude, de grandeur etc., qu'elle soutient avec toutes les autres choses, est impossible. C'est pourquoi, dans les raisonnements, nous lui substituons toujours une chose incomplète ou plus ou moins complétée.

Cette distinction est essentielle, parce que *la possibilité apparaît seulement dans la catégorie des choses incomplètes.* On voit en effet que, s'agissant d'un objet concret ou d'une chose complète, une proposition qui s'y rapporte est ou bien juste ou bien fausse ; il n'y a pas place pour un moyen terme. La notion « c'est possible » apparaît seulement quand il s'agit de choses incomplètes. Ainsi, il y a possibilité qu'un triangle soit rectangle. Mais cette possibilité ne saurait se rapporter à un triangle que je vois devant moi ; pour cette chose complète, la proposition « Il est rectangle » est nécessairement ou juste ou fausse, une troisième alternative est impensable. La *possibilité* d'être rectangle n'existe que pour le triangle en soi, pour la notion de triangle.

131. — Au point de vue philosophique, l'étude de la notion de possibilité présente un grand intérêt. Elle diffère très essentielle-

ment de l'étude des propriétés physiques et des phénomènes matériels. Ici, en effet, la réalité extérieure constitue la base, nos sens font l'office d'intermédiaires et fournissent à l'intelligence les matériaux de la connaissance. Les couleurs se voient, les sons s'entendent etc. ; les résidus qu'ils laissent dans la mémoire forment le *substratum* du travail intellectuel d'abstraction qui aboutit à la notion de couleur, à celle de son etc. Rien de semblable pour la notion de possibilité, rien de palpable ni de matériel qui se présente immédiatement à la pensée. La réalité matérielle est même trop étroite, puisqu'en général, des objets ne peuvent pas donner lieu à des possibilités. L'exemple suivant, imaginé par Mr Meinong, fait voir de quelle source jaillissent les possibilités bien fondées.

Imaginons une société fermée, par exemple l'assemblée d'une corporation ; elle a délibéré à huis-clos, en séance secrète ; mais malgré l'exclusion rigoureuse de toute personne étrangère à la société, certains détails sur les délibérations sont parvenus à la connaissance publique. Sur chaque membre individuellement planera le soupçon de s'être rendu coupable d'indiscrétion ; le soupçon se répartira par parts égales sur les sociétaires et chaque membre indistinctement y participera. Or, qui dit soupçon dit possibilité ; ici, on voit bien clairement qu'*un fait* est à la base de la possibilité, le fait que des nouvelles ont été répandues illicitement dans le public. De cet exemple particulier, on arrive par voie d'abstraction à la proposition générale suivante.

Définition 2. — Soit un ensemble M de choses quelconques, semblables ou disparates, dites *éléments* de l'ensemble M. S'il est connu que l'un de ces éléments, sans qu'on sache lequel, possède effectivement le caractère X, il en résultera pour chaque élément de M la possibilité d'être porteur du dit caractère. C'est en cela que consiste *le principe de la participation*. A l'égard de ce principe, les n éléments de la collectivité M sont équivalents entre eux, aussi bien celui qui est effectivement porteur de la qualité X que les $(n - 1)$ autres. Il est fait abstraction de toutes leurs propriétés individuelles, sauf de celle d'appartenir à l'ensemble M.

Définition 3. — Les ensembles les plus usuels que l'on considère dans le calcul des probabilités sont *les ensembles de collocation*, c'est-à-dire, comme l'indique l'étymologie, ceux dont les éléments se trouvent réunis dans un même lieu, par exemple les boules d'une même urne, les cartes d'un même jeu, les deux côtés d'une pièce de monnaie, les 6 faces d'un dé etc.

132. — *Le principe de la participation* consiste en ceci : le caractère X, au lieu de rester confiné à celui des n éléments qui le possède effectivement, se transmet en quelque sorte à tous les autres et se répartit également entre eux sous forme de possibilité. Chaque élément en reçoit la $n^{ième}$ partie. Transmission et répartition uniforme, telles sont les deux idées fondamentales dans ce principe. Si l'on fait correspondre un nombre au caractère envisagé X, mettons le nombre 1 qui se présente le plus naturellement, il correspondra à chacun des n éléments le nombre $1 : n$.

Si le caractère en question appartient à plus d'un élément, par exemple à f éléments de l'ensemble envisagé, la transmission et la répartition uniforme se font f fois, indépendamment les unes des autres. Ces participations s'additionnent, de sorte qu'à chaque élément de notre collectivité correspondra le nombre $f : n$.

Cette conclusion présuppose l'équivalence rigoureuse des éléments ; c'est sur elle, et même sur elle seule que se fonde la mesurabilité. Or, cette équivalence réside dans le fait que les éléments sont envisagés uniquement en leur qualité d'éléments, c'est-à-dire comme des choses appartenant à l'ensemble considéré. On fait abstraction, par définition et par hypothèse, de toutes autres propriétés individuelles. Cela exclut aussi tout facteur subjectif et donne à la possibilité mesurée un caractère d'objectivité.

133. — Les ensembles de collocation ne forment qu'une partie des matières donnant lieu à des possibilités mesurables. En approfondissant ce sillon, on est amené à faire une distinction entre les deux notions de « réalité » et « d'existence ».

Définition 4. — La *réalité* est un attribut des choses concrètes seulement. L'*existence* se rapporte aussi à des choses abstraites, concevables à priori.

Exemples. — La notion de triangle a de l'existence, non de la réalité. Un triangle rectangle plan avec trois angles droit n'a ni existence ni réalité. Un triangle rectangle est possible et a de l'existence. Quand il est tout tracé devant moi sur le papier, ou fabriqué mécaniquement, découpé en bois par exemple, il a en outre de la réalité. La notion d'existence ainsi entendue s'applique donc aussi à des choses empiriquement connaissables

La notion de probabilité n'est pas confinée aux choses ayant réalité ; elle est applicable encore dans des problèmes où il s'agit de spéculations purement abstraites.

On voit aussi que la fraction $f : n$ n'est pas la probabilité mathématique elle-même, mais la mesure de cette probabilité. Déjà Laplace distingue entre ces deux notions, du moins implicitement ; chez Poisson, cette distinction est tout-à-fait explicite, lorsqu'il dit : « *La mesure* de la probabilité d'un événement aléatoire est le rapport du nombre de cas favorables à cet événement, au nombre total de cas favorables ou contraires, pourvu qu'ils aient tous la même chance ».

Remarque. — D'après le principe de la participation, la possibilité a sa source dans un fait ; par exemple, l'une des n boules d'une urne, on ne sait pas laquelle, est effectivement rouge. Ce fait produit la possibilité chez chacun des n éléments. Ces n possibilités sont égales entre elles. Il correspond alors à chacune le nombre $1/n$, si l'on fait correspondre au fait lui-même le nombre 1. L'égalité de ces n possibilités est basée uniquement sur l'appartenance à l'ensemble, puisqu'on fait abstraction de toutes autres propriétés individuelles des éléments. Si l'on voulait étayer cette égalité sur un principe philosophique, ce ne pourrait être que sur le principe de l'absence de raison justifiante (v. n° 18 et n° 125, remarque), car il n'y a pas d'autre raison que l'appartenance à l'ensemble M. Il s'agit ici de *l'inexistence logique d'une différence entre les éléments.*

Cette interprétation du dit principe lui enlève tout caractère subjectif. En effet, on ne s'en rapporte plus à l'état des connaissances de celui qui juge, état qui pourrait varier d'une personne à l'autre, mais à quelque chose d'objectif valable pour l'intelligence même

la plus éclairée. Ce n'est plus un aveu d'ignorance : « Il me manque une raison de faire une différence entre les éléments ». Non ! car cette raison manquerait à chacun, même à un esprit omniscient, parcequ'elle n'existe pas en réalité ; on retient, en effet, uniquement la propriété de faire partie du même ensemble M et l'on fait abstraction de tout autre caractère.

II. — *Probabilité mathématique et théorie des ensembles.*

134. — Les développements qui précèdent vont permettre de définir la probabilité mathématique d'une manière complète et définitive. Par *notion générale*, entendons un caractère ou un groupe de caractères commun à plusieurs choses, le terme de « chose » étant pris dans le sens le plus large possible, comme au numéro 130. Un caractère peut d'ailleurs être mesurable ou non (v. n° 179).

Définition 5. — Une pluralité de choses constitue *une classe* ou *un ensemble*, dont ces choses sont dites *les éléments*. Il peut arriver qu'une classe ne contienne pas d'élément du tout ; elle est alors dite *vide*.

Définition 6. — Les ensembles les plus importants sont les ensembles *bien définis*, c'est-à-dire où l'on peut décider si une chose quelconque donnée fait partie de l'ensemble ou non. Il ne sera question dans la suite que d'ensembles bien définis. Si le caractère commun à tous les éléments, la notion servant à définir l'ensemble, est mesurable ou peut s'exprimer par des nombres, j'appellerai l'ensemble un *syllepte* ou un *collectif* (v. n° 179).

Définition 7. — La totalité des caractères communs à tous les éléments d'un ensemble, ou la totalité des qualités qui leur sont communes, est une *notion*. La définition de cette notion, c'est la proposition indiquant par quel nom on désigne ce groupe de caractères ; toute définition est ainsi une explication d'un nom.

Les notions elles-mêmes peuvent à leur tour être envisagées comme des unités et constituer les éléments d'un nouvel ensemble ou d'une nouvelle classe, et ainsi de suite. On peut s'élever de

cette façon à la conception d'éléments et d'ensembles de plus en plus abstraits.

Une notion délimite un ensemble de choses dont chacune possède encore d'autres caractères que ceux indiqués dans la notion. Toute chose, cela va sans dire, porte nécessairement avec elle tous les caractères et toutes les qualités qui lui sont propres, toutes les relations, en nombre infini, qu'elle soutient avec les autres choses. Mais quand on l'envisage comme élément d'un ensemble, on ne tient compte que d'une partie de ces caractères et propriétés, on laisse de côté tous les autres ; c'est cette opération intellectuelle qu'on nomme abstraction.

Remarques. — La formation d'ensembles est un processus idéal qui ne change rien aux choses elles-mêmes. S'il peut être important pour moi et pour mes actions futures de diviser les plantes en deux ensembles, les cryptogames et les phanérogames, cela ne touche en rien les plantes elles-mêmes. Observons enfin que la formation d'ensembles est une opération tout-à-fait arbitraire. Elle n'est déterminée que par les buts que l'on se propose d'atteindre. Rien n'empêche de réunir les caractères les plus disparates pour en constituer un ensemble.

Ces préliminaires posés, passons à la définition de la probabilité mathématique.

135. — Soit M un ensemble défini par le caractère ou le groupe de caractères A, de sorte que les éléments de M sont tous porteurs des qualités A. Soit T un sous-ensemble de M, tel que tous les éléments de T possèdent, en plus du caractère A servant à définir M, encore le caractère ou le groupe de caractères B. Trois cas sont possibles : cet ensemble T peut comprendre tous les éléments de M ou quelques-uns seulement ou aucun. Dans ce troisième cas, T est un ensemble vide ; dans le second, T est dit *un diviseur propre* de l'ensemble M, et dans le premier cas, T est dit *diviseur impropre* de M.

Les caractères A et B sont *indépendants* les uns des autres. C'est d'une indépendance au point de vue logique qu'il s'agit ici. S'il y avait dépendance logique entre les qualités A et B, l'existence

de A entraînerait celle de B et B devrait être, comme A, une propriété de tous les éléments de M. Nous supposons donc qu'il n'en soit pas ainsi. Les éléments de M sont alors « *non-déterminés* par rapport aux qualités B » ; en d'autres termes : de ce qu'une chose fait partie de M, on ne peut pas conclure qu'elle possède nécessairement le caractère B ; elle peut l'avoir ou ne pas l'avoir. Chaque élément de M, pris individuellement et considéré comme représentant la notion A, est non-déterminé par rapport à toutes les qualités et propriétés qui ne sont pas comprises dans la notion A ou impliquées par elle.

Il va de soi que le caractère d'être « non-déterminé » ne peut appartenir à un élément que par rapport à une classe ou à un ensemble ; car, ainsi que nous le faisions déjà remarquer au numéro précédent, une chose porte nécessairement avec elle tous ses caractères et toutes les qualités qui lui sont propres. Le caractère B n'étant pas compris dans la notion A qui sert à définir l'ensemble M, ni impliqué par A, l'apparition de B chez un élément de M est fortuite ; c'est un *hasard dans le sens logique* ; c'est de lui qu'il s'agit dans le calcul des probabilités.

Définition 8. — Tout caractère distinctif non contenu dans la définition et non impliqué par elle, mais appartenant en propre à un élément, est jugé lui appartenir *fortuitement* ou *par hasard*. Cette importante notion de *hasard dans le sens logique* conduit à poser la

Définition 9. — On appelle *cas également possibles* ou *cas équipossibles*, des éléments faisant partie d'un même ensemble.

Avant de passer à la probabilité mathématique, rappelons quelques définitions de la théorie des ensembles.

Définition 10. — Soient deux collectivités, M_1 et M_2. On dit que M_2 est *une image de* M_1, si à tout élément de M_1 correspond un élément de M_2 et un seul. Le terme d'image prend ici un sens technique très différent de l'acception courante du mot « image ».

Définition 11. — On dit que M_1 et M_2 se correspondent *parfaitement*, s'il y a entre eux une correspondance uni-univoque ; en

d'autres termes, si à tout élément de M_1 correspond un élément bien déterminé de M_2 et un seul, et que réciproquement à tout élément de M_2 corresponde un élément bien déterminé de M_1 et un seul. A des éléments différents correspondent alors des images différentes.

Définition 12. — Deux ensembles ont *même puissance*, si l'on peut établir entre eux une correspondance parfaite.

136. — Dans la théorie des ensembles, on fait correspondre à tout ensemble un nombre réel bien déterminé, dit *la mesure* de l'ensemble.

Définition 13. — Quand un ensemble ne contient qu'un nombre fini d'éléments, sa puissance et sa mesure sont données par le nombre de ses éléments.

Des définitions 9 et 13 découle ce *théorème* : deux sous-ensembles d'un même ensemble M sont équipossibles, s'ils ont même mesure.

Notation. — Nous désignerons la mesure d'un ensemble M par le symbole $\overline{M}$ ou $Me(M)$, prononcé « mesure de M ».

Soit comme ci-dessus l'ensemble M défini par le caractère ou le groupe de caractères A. Envisageons à nouveau le sous-ensemble T caractérisé par la simultanéité des qualités A et B. Supposons d'abord l'ensemble total M fini. Il s'ensuivra que T l'est également et que les mesures de ces deux ensembles, $Me(T)$ et $Me(M)$, sont données par les nombres respectifs de leurs éléments. Formons le rapport de ces mesures et désignons-le par p.

$$p \equiv \frac{\overline{T}}{\overline{M}} = \frac{Me(T)}{Me(M)} \qquad \text{(VI 1)}$$

Ce rapport est un nombre réel compris entre 0 et 1, limites incluses si l'on admet des classes vides et des diviseurs impropres.

Définition 14. — Nous dirons que le rapport défini par (VI 1) est « *la probabilité mathématique* qu'un élément de M fasse partie du sous-ensemble T » ou encore, en tenant compte de la

définition des deux ensembles, *la probabilité mathématique du caractère* B, *quand le caractère* A *existe*. Si le caractère *B* ne peut appartenir qu'à des éléments de *M*, on peut, sans crainte de confusion, parler de *la probabilité mathématique de B*, tout court ([69a]).

Cette définition suppose les ensembles *M* et *T* bien définis ; en particulier, l'ensemble total *M* doit être bien délimité. En effet, si l'on considérait l'ensemble de toutes les choses existantes, la probabilité mathématique du caractère *B* serait toujours nulle, car $\overline{T} : \infty = 0$.

Quand on parle de la probabilité mathématique de *B*, tout court, l'ensemble total *M* est sous-entendu, quoique parfaitement délimité. Par exemple, la propriété d'être nombre premier ne peut appartenir qu'à un nombre entier ; donc, « la probabilité mathématique d'un nombre premier » ne peut signifier que la probabilité pour qu'un nombre *entier* soit premier. De même, « la probabilité mathématique d'un choc de deux molécules » ne peut se rapporter qu'à un ensemble de molécules tel que ceux qui interviennent dans la théorie cinétique des gaz. En dehors des domaines en question, les choses dont il s'agit ont une probabilité nulle.

Si le caractère *B* appartient à des éléments de plusieurs ensembles, M_2, M_3, ..., il faut indiquer à laquelle de ces collectivités la probabilité mathématique doit se rapporter ; car, répétons-le, la probabilité mathématique ne se rapporte pas à l'existence en soi du caractère *B*, mais à la coexistence des caractères *A* et *B*.

137. — Des définitions ci-dessus posées, on peut tirer les conséquences suivantes. 1) *Le hasard dans le sens logique a toujours un caractère de relativité*. D'ailleurs, en analysant complètement la notion de hasard, on voit qu'il ne saurait y avoir de hasard absolu ou extra-causal, si l'on entend par ce terme l'absence de cause. Il est vrai que plusieurs auteurs, en particulier Em. Czuber ([70]),

(69 *a*) Voir F. M. Urban, *l. cit.* (77), p. 57.

(70) Emanuel Czuber. *Die philosophischen Grundlagen der Wahrscheinlichkeitsrechnung*, Leipzig und Berlin, 1923 (collection « Wissenschaft und Hypothese », n° 24), p. 179.

adoptent le point de vue que le Calcul des probabilités fait théoriquement abstraction de toute relation de cause à effet. Selon ces penseurs, il faut, dans le calcul des probabilités, poser d'emblée l'hypothèse d'événements tout-à-fait indépendants, sans connexion quelconque, ni directe ni indirecte ; à eux seulement, les théorèmes stocastiques seraient rigoureusement applicables. En adoptant la définition 14, on n'est pas obligé de recourir à cette fiction.

2) *La probabilité mathématique*, étant le rapport des mesures de deux ensembles, *a de la réalité* dans le sens du numéro 133, définition 4. C'est une grandeur objective indépendante de l'état de mes connaissances ; elle existe, même si mon attention ne la perçoit pas. Notre définition 14 a l'avantage de ne présupposer aucune relation entre un sujet pensant et les grandeurs examinées ; c'est une probabilité sans conjecture ni présomption.

Définition 15. — Le choix fortuit. La probabilité mathématique p définie par (VI 1) se rapporte à la présence du caractère ou du groupe de caractères B, chez un élément quelconque de l'ensemble M défini par le caractère ou le groupe de caractères A. On *choisit* un élément, c'est-à-dire qu'on le sort ou qu'on l'envisage ou qu'on le considère, en tenant compte uniquement de son appartenance à l'ensemble M et faisant abstraction de toute autre propriété ou caractère ou relation. Quant à savoir si un procédé physique ou mécanique déterminé réalise effectivement cette idée de « tirage au hasard » ou *choix fortuit*, c'est une autre question ; elle exige dans chaque cas particulier un examen détaillé. Remarquons seulement que dans les règles de tous les jeux de hasard, certaines prescriptions ont pour but, et le plus souvent aussi pour effet, d'amener cette équivalence complète des éléments, en ne laissant subsister, au point de vue du choix ou résultat final, que leur appartenance à l'ensemble ; tels sont : le battage des cartes avant leur distribution, le brassage prolongé des boules ou des numéros dans l'urne, les mouvements imprimés au dé dans le cornet etc.

138. — La définition de la probabilité mathématique ci-dessus posée, en utilisant les notions de la théorie des ensembles, est un développement des idées de J.-J. Fourier, de A.-A. Cournot, de

Jean de Kries et de F.-A. Lange. Fourier, Cournot et de Kries ont recours à la notion un peu vague de domaine ou d'étendue ou de « spaciom comportaire », comme nous disons. Vu qu'il est souvent impossible de préciser numériquement ce qu'il faut entendre par là, ces auteurs n'évitent pas quelque indétermination. Seule la théorie des ensembles permet de donner à ces notions toute la précision dont elles sont susceptibles, et ni Fourier ni Cournot ne l'avaient à leur disposition.

En démontrant que la notion de probabilité mathématique est de nature purement logique, on démontre du même coup l'inanité de l'interprétation empiriste exposée au numéro 112. Il en est de même des essais d'identifier la probabilité mathématique avec ce que j'appelais les nombres analogiques (v. n° 112, définition 2) ou avec la fréquence relative ou avec la fréquence dite moyenne, et de restreindre ainsi le domaine du Calcul des probabilités aux événements où l'on peut prouver la constance des fréquences relatives. Si les applications de la théorie des chances sont limitées à ces domaines-là, le Calcul des probabilités, comme partie des mathématiques pures, a un domaine plus vaste. L'idée empiriste a ceci de juste qu'elle rappelle l'origine historique de ce Calcul. Comme presque toutes les théories mathématiques, il s'est développé en connexion avec l'expérience et parcequ'on cherchait à expliquer certains phénomènes.

Probabilité et chance

139. — Lorsqu'on tente de déterminer la valeur numérique d'une probabilité mathématique par l'observation de phénomènes réels, il faut distinguer entre la valeur que cette grandeur a effectivement et la valeur que donnent nos moyens d'investigation toujours limités et imparfaits. La valeur objective, définie par (VI 1) comme rapport des mesures de deux ensembles, existe indépendamment de notre perception et de nos mesures ; mais elle nous est souvent inconnue, comme beaucoup de constantes physiques. Nous devons dès lors nous contenter d'une valeur plus ou moins approximative qui sera toujours susceptible de corrections et oscillera entre certaines limites d'indétermination. Déjà S.-D. Pois-

son, dans un ouvrage qui fit sensation ([71]), distinguait entre ces deux valeurs de la probabilité mathématique, bien que ses définitions n'aient pas toujours la netteté désirable. Il appelle la première valeur *la chance* ou *la probabilité abstraite*, et l'autre *la probabilité subjective*. Son idée fut reprise par Cournot qui exposa ses vues sur ce sujet dans une lettre à Poisson ([72]).

Plusieurs auteurs modernes, notamment Em. Czuber, font une distinction analogue. Elle n'a rien à voir avec celle établie par Henri Poincaré dans *La science et l'hypothèse*, p. 218. Là, Poincaré confond, du moins dans la terminologie, d'une part « probabilité subjective » avec probabilité à priori, d'autre part « probabilité objective » avec probabilité à postériori. Les termes de probabilité à priori et probabilité à postériori sont beaucoup plus adéquats aux notions qu'ils expriment et d'ailleurs presqu'universellement adoptés, de sorte qu'il n'y a pas lieu de suivre la terminologie de Poincaré, due peut-être à une inadvertance.

Définition 16. — Avec Poisson et la plupart des auteurs, nous réserverons le terme de *chance d'un événement* pour désigner une grandeur objective, savoir la véritable valeur, connue ou inconnue, de la probabilité mathématique de cet événement.

Il suit de la définition 14 (v. n° 136) que, pour déterminer la probabilité mathématique $p(E)$ d'un phénomène E, il faut certaines connaissances. Elles doivent rendre possible la détermination du rapport des mesures des deux ensembles M et T; si ces connaissances-là font défaut, il devient impossible de déterminer $p(E)$. C'est le cas en particulier quand il s'agit de choses que l'on ne peut pas soumettre avec succès au travail d'abstraction conduisant à la constitution d'un ensemble, ou lorsqu'il manque l'ensemble M auquel la probabilité mathématique doit se rapporter. Par exemple, « la probabilité mathématique qu'une proposition soit juste » est un assemblage de mots dénué de sens; car toute proposition, étant une chose complète dans le sens de la définition 1 du numéro 130,

(71) Voir la note (3).

(72) La réponse de Poisson, datée du 26 janvier 1836, est reproduite p. VI de la Préface du livre de Cournot cité dans la note (31). Voir aussi F. Faure, *Les idées de Cournot sur la statistique*, dans la Revue de Métaphysique et de Morale, Paris, 1905, t. XIII, p. 404.

porte avec elle toutes ses qualités et propriétés, entre autres celle d'être juste ou d'être fausse. Tout au plus pourrait-on parler, dans ce cas, de probabilité philosophique ou de probabilité psychologique, lorsqu'il s'agit de considérations philosophiques ou de circonstances où interviennent des facteurs psychologiques.

III. — *Les paradoxes du calcul des probabilités.*

140. — Les définitions posées au numéro 136 permettent d'expliquer certains paradoxes de la théorie des chances. Ils proviennent de l'indétermination des ensembles M et T qui entrent dans la définition. En général, cette indétermination est due à ce que la question posée laisse subsister des doutes sur ce qu'il faut entendre par M ou par T. C'est une particularité de beaucoup de problèmes se rapportant aux probabilités dites géométriques, où il s'agit d'ensembles infinis. C'est que l'infini n'est pas un nombre déterminé ; on ne doit pas sans explication l'introduire dans les raisonnements. Il arrive donc que le problème posé ne délimite pas univoquement l'ensemble M ; il en laisse même la définition précise à l'arbitraire du lecteur ; de là, une multiplicité de solutions différentes, c'est tout naturel (v. n° 147).

L'exemple le plus célèbre est connu dans la littérature sous le nom de *paradoxe de Bertrand*. Voici ce problème. « *On trace au hasard* une corde dans un cercle donné de rayon r. Quelle est la probabilité mathématique p pour que cette corde soit plus grande que le côté s du triangle équilatéral inscrit ? » La particularité de ce problème consiste en ceci : l'ensemble des cordes « tracées au hasard », donc l'ensemble M de la définition (VI 1), n'est pas univoquement déterminé. La précision illusoire des mots fait naître des contradictions. On peut en effet interpréter de plusieurs façons les mots « tracer une corde au hasard », et l'énoncé du problème ne dit pas laquelle de ces interprétations doit être entendue. En voici les plus simples.

1) On peut se donner la direction de la corde AB ; alors la posi-

tion du point P où elle coupe le diamètre d perpendiculaire à sa direction doit être réglée par le hasard.

2) On peut se donner l'une des extrémités de la corde AB, par exemple A ; alors sa direction doit être réglée par le hasard.

3) Pour fixer la position d'une corde, il suffit de se donner son point milieu C. On peut donc choisir au hasard l'un des points du cercle comme point milieu C de la corde ACB.

4) On peut se donner l'une des extrémités de la corde, par exemple A, et choisir au hasard dans le cercle un second point, P, qui fixera la position de la corde APB.

5) On peut choisir au hasard, sur la circonférence donnée, les deux points A et B, extrémités de la corde en question.

6) On peut choisir au hasard deux points de la surface du cercle, P_1 et P_2 ; ils fixent la position de la corde AP_1P_2B.

On pourrait allonger encore la liste des possibilités, en considérant par exemple les sécantes issues d'un point extérieur au cercle etc. On voit que la pluricité de la solution provient de ce que, pour pouvoir résoudre le problème, on est obligé de choisir une donnée qui n'est pas impliquée par l'énoncé du problème. Le même énoncé comprend donc plusieurs problèmes différents, ce qu'Ernest Cesàro semble avoir été le premier à remarquer ([73]).

141. — De chacune des hypothèses mentionnées au numéro 140, on peut déduire une solution juste. Les voici.

Ad 1). La direction de la corde AB étant choisie, arbitrairement d'ailleurs, pour que cette corde soit plus longue que le côté s du triangle équilatéral inscrit, elle doit être aussi plus rapprochée du centre. Le point P où elle coupe le diamètre d perpendiculaire à sa direction doit donc se trouver plus près du centre, savoir sur un segment de longueur r placé symétriquement par rapport au centre, car la distance au centre, d'un côté du triangle équilatéral inscrit, est égale à la moitié du rayon. Or, l'ensemble M de la définition (VI 1) comprend ici toutes les cordes parallèles qui coupent le diamètre d perpendiculaire à leur direction ; donc $Me(M) = 2r$. Le sous-ensemble T comprend celles de ces cordes dont le point d'in-

(73) Ernesto Cesàro, *Considerazioni sul concetto di probabilità* Periodico di Matematica, t. VI, 1891, p. 1-13, 40-62.

tersection avec d est sur un segment de longueur r; donc $Me(T) = r$. Il en découle que la probabilité mathématique demandée est $p = r : 2r = 1/2 = 0{,}50$.

Ad 2). L'extrémité A de la corde étant choisie au hasard sur la circonférence, la tangente en A et les deux côtés du triangle équilatéral inscrit ayant A pour sommet forment trois angles égaux chacun à 60^0. Pour que la corde soit plus longue que s, sa direction doit être intérieure à l'angle que forment les deux côtés. L'ensemble M comprend ici toutes les cordes issues de A ; leurs directions remplissent un angle de 180^0, de sorte que $Me(M) = \pi$. Le sous-ensemble T comprend les cordes issues de A et intérieures à un angle de 60^0, de sorte que $Me(T) = \frac{\pi}{3}$. La probabilité mathématique cherchée est donc $p = \pi/3 : \pi = \frac{1}{3} = 0{,}3333\ldots\ldots$

Ad 3). L'ensemble M comprend ici toutes les cordes dont les points milieu couvrent la surface du cercle donné, de sorte que $Me(M) = r^2\pi$. Pour qu'une corde satisfasse à la condition de l'énoncé, son point milieu C doit être plus près du centre du cercle que ne l'est le point milieu du côté s ; donc C doit être intérieur à un cercle concentrique au cercle donné et de rayon moitié, d'où $Me(T) = \frac{1}{4} r^2\pi$. La probabilité mathématique cherchée est donc $p = \frac{1}{4} r^2\pi : (r^2\pi) = \frac{1}{4} = 0{,}25$.

Ad 4). L'extrémité A étant choisie au hasard sur la circonférence, supposons tracés les deux côtés AF et AG du triangle équilatéral inscrit correspondant AFG. Si le point P arbitrairement choisi tombe à l'intérieur du secteur GAF, on aura un cas favorable ; donc, le sous-ensemble T, comprenant tous les points intérieur à ce secteur, a pour mesure $r^2\left(\frac{\pi}{3} + \frac{1}{2} \cdot \sqrt{3}\right)$. D'autre part, $Me(M) = r^2\pi$. La probabilité mathématique cherchée est donc

$$p = r^2\left(\frac{\pi}{3} + \frac{1}{2}\sqrt{3}\right) : (r^2\pi) = \frac{1}{3} + \frac{\sqrt{3}}{2\pi} = 0{,}608\ 997\ldots.$$

Ad 5). Cette hypothèse revient à celle du cas 2). En effet, l'extrémité A de la corde étant choisie arbitrairement sur la circon-

férence donnée, tout rayon issu de A détermine sur elle un second point, B, et vice-versa. A cause de cette correspondance parfaite, on retrouve comme probabilité mathématique cherchée

$$p = 1/3 = 0,333\,33\ldots\ldots$$

Ad 6). Si le premier point choisi au hasard, P_1, tombe dans le cercle concentrique au cercle donné et de rayon moitié, ce qui a la probabilité mathématique $1/4$ (v. chiffre 3), toute corde tracée par P_1 satisfera aux conditions de l'énoncé. Mais si P_1 est extérieur au dit cercle, la corde P_1P_2 ne pourra satisfaire à ces conditions que si P_2 tombe à l'intérieur de deux secteurs dont le calcul infinitésimal permet de déterminer l'aire. On trouve ([74]), pour la probabilité mathématique cherchée,

$$p = \frac{4\pi + 9\sqrt{3}}{12\pi} = 0,746\,830\ldots\ldots$$

Le résultat donné par Mr Lämmel dans sa thèse de doctorat ([75]), $p = (5\pi + 9\sqrt{3}) : (16\pi) = 0,622\ldots\ldots$, est inexact.

142. Joseph Bertrand ne donne que les trois premières solutions et résume comme suit son point de vue : « Entre ces trois réponses, quelle est la véritable ? Aucune des trois n'est fausse, aucune n'est exacte, la question est mal posée ». De ces trois affirmations de Bertrand, la seconde est inacceptable. En effet, si les réponses sont différentes, c'est parcequ'elles reposent sur des données différentes. La solution d'un problème mathématique est juste si elle est déduite des données par des raisonnements logiques impeccables. Donc, chacune des 6 réponses qu'on vient de faire au problème de Bertrand est exacte et juste au sens mathématique. Il en serait autrement s'il s'agissait d'un procédé empirique, d'un mécanisme physique qui effectivement « tracerait au hasard » des cordes dans un cercle donné. Exemple : on lance un

(74) Voir p. 118 de l'ouvrage d'Em. Czuber cité dans la note (46).

(75) Rudolph Laemmel. *Untersuchungen über die Ermittelung von Wahrscheinlichkeiten*, Inauguraldissertation, Zürich, 1904, p. 29-32, 38-39.

disque circulaire de diamètre $d = 2r$ sur une table recouverte de droites systématiquement tracées, mettons d'un ou de deux faisceaux de parallèles équidistantes, l'équidistance étant supérieure à d. L'analyse de ce procédé mécanique, ou de tout autre que l'on appliquerait, déterminerait alors l'hypothèse dont il faudrait partir pour arriver à la solution du problème de Bertrand, c'est-à-dire déterminerait ce qu'il faudrait entendre par « cordes tracées au hasard », soit les ensembles M et T. Et seule la solution déduite de cette hypothèse-là serait la vérité empirique. Les autres solutions, sans cesser d'être les réponses exactes à des problèmes analogues, ne seraient plus vérifiées par les résultats expérimentaux du procédé en question. Il n'y a donc aucun paradoxe dans le fait que le problème de Bertrand comporte diverses solutions, pas plus qu'il n'y a paradoxe dans le fait qu'on trouve des résultats différents pour l'aire d'un terrain rectangulaire, quand on admet pour les côtés des longueurs différentes résultant de plusieurs procédés de mesure.

IV. — *Probabilité mathématique et ignorance.*

143. — Nous avons déjà eu l'occasion d'observer que de l'ignorance absolue, on ne saurait tirer des connaissances positives, même à l'aide des plus élégantes formules mathématiques (v. n° 125, remarque). Cela devrait aller de soi, semble-t-il ; et pourtant, la plupart des auteurs sont de l'avis que le manque de connaissances est une condition nécessaire pour qu'on puisse appliquer le calcul des probabilités. « Le problème des probabilités existe seulement là où il y a ignorance », disent-ils ; il faut « une égale répartition de l'ignorance ». Ou encore : « La théorie de la probabilité s'occupe de la quantité du savoir ; un événement est seulement probable, lorsque la connaissance que nous en avons est mélangée d'ignorance et qu'il faut un calcul exact pour discerner combien nous en savons et combien nous en ignorons » ([76]). On pourrait

(76) W. Stanley Jevons, *The principles of science.* A treatise on Logic and scientific method. London, 1892, p. 199.

citer à foison des passages qui abondent dans ce sens (77), jusqu'à cette définition du calcul des probabilités : « c'est la science de systématiser le non-savoir » (78).

Beaucoup d'auteurs pensent que les lois régissant les phénomènes doivent être partiellement inconnues. Nous avons démontré que le contraire est vrai. L'ignorance doit porter tout au plus sur les circonstances ontologiques ou conditions initiales (v. n° 117 et n° 118, *b*). Or, il suit de notre dernière définition que même cela n'est point nécessaire. Nous dirions plutôt que le calcul des probabilités est une méthode permettant d'éviter de baser des conjectures sur l'ignorance.

Henri Poincaré a fait remarquer une fois avec beaucoup d'esprit : « Vous me demandez de vous prédire les phénomènes qui vont se produire. Si par malheur je connaissais les lois de ces phénomènes, je ne pourrais y arriver que par des calculs inextricables et je devrais renoncer à vous répondre. Mais comme j'ai la chance de les ignorer, je vais vous répondre tout de suite ; et, ce qu'il y a de plus extraordinaire, c'est que ma réponse sera juste » (79). On a souvent, et avec raison, combattu ce point de vue du on-sait-d'autant-plus-qu'on-ignore-davantage (80). La meilleure réfutation est donnée par Poincaré lui-même, quand il écrit en connaisseur du calcul des probabilités et non en vulgarisateur faisant de l'esprit. Dans son ouvrage sur le calcul des probabilités (81), il traite d'une manière aussi intéressante qu'originale le jeu de la roulette et arrive au résultat que le problème mathématique serait insoluble si l'on ne savait absolument rien de deux fonctions, f et φ, qu'il

(77) Voir Friedrich M. Urban, *Grundlagen der Wahrscheinlichkeitsrechnung und der Theorie der Beobachtungsfehler*, Leipzig et Berlin 1923, p. 73 et suivantes.

(78) Voir la note (28), p. 26, et la fin du numéro 18.

(79) Henri Poincaré, *Science et méthode*, Paris, 1909, voir le chapitre intitulé « Le hasard ».

(80) Par exemple Félix-A. Le Dantec, *De l'homme à la science*, Paris, 1907, chap. 12. *Le chaos et l'harmonie universelle*, Paris, 1911.

(81) Henri Poincaré, *Calcul des probabilités* ; rédaction de A. Quiquet, 2e édition, revue et augmentée par l'auteur, Paris, 1912. Nouveau tirage, Paris, 1923, p. 148-150.

introduit dans son analyse ; ici, on doit savoir que f a une dérivée non infinie ; c'est fort peu de choses des propriétés de f, mais cela suffit ; et Poincaré écrit lui-même : « Si on ne savait rien du tout sur φ ou sur f, on ne pourrait rien calculer ; c'est parceque l'on sait quelque chose que l'on peut entreprendre le calcul ». Il en est ainsi dans tous les cas où l'on applique avec succès les théorèmes de la stocastique.

144. — Nous allons plus loin. Nous disons que même si notre connaissance du domaine ou du phénomène en question était complète, que nous sachions tout ce qui se rapporte aux circonstances ontologiques aussi bien que nomologiques, il pourrait y avoir lieu d'appliquer le calcul des probabilités. On s'en rend aisément compte à priori. En effet, les renseignements fournis par le Calcul des probabilités sur certains événements aléatoires ne cesseront pas d'être vrais le jour où ces phénomènes seront mieux connus. Le directeur d'une Compagnie d'assurance sur la vie humaine ignore totalement quand mourra chacun de ses assurés. Mais il compte sur la loi des grands nombres pour prévoir à peu près le nombre des décès à venir. Et il ne se trompe pas, fait remarquer Poincaré, puisqu'il distribue des dividendes à ses actionnaires. Ces dividendes ne s'évanouiraient pas si un médecin très perspicace et très indiscret, doué d'un savoir surhumain, venait, une fois les polices signées, renseigner le directeur sur les chances de vie des assurés. « Ce médecin dissiperait l'ignorance du directeur », dit Poincaré, « mais il n'aurait aucune influence sur les dividendes qui ne sont évidemment pas un produit de cette ignorance. »

La définition formulée au numéro 136 a entre autres avantages celui de montrer immédiatement les conditions nécessaires et suffisantes pour qu'une probabilité mathématique puisse être établie : les connaissances doivent être assez étendues pour permettre de calculer le rapport de la mesure de deux ensembles. Cette définition ne postule donc aucune ignorance.

145. — Le caractère ou le groupe de caractères A qui délimite l'ensemble M, et le caractère ou le groupe de caractères B qui en détermine le sous-ensemble T, nous pouvons les fixer plus ou

moins à notre gré et selon notre convenance. Il n'est point nécessaire d'utiliser à cet effet toutes les connaissances que l'on pourrait avoir sur la matière envisagée. Gauss avait raison quand il disait, en matière d'introduction, que le Calcul des probabilités a pour but de nous renseigner, mais seulement dans les cas où, en dehors des chiffres fournis par l'observation, « nous ne *savons* rien ou *ne voulons pas* tenir compte d'autre chose ». Il y a même de nombreux cas où l'on fait intentionnellement abstraction d'une partie des informations que l'on possède et où néanmoins l'investigation ainsi conduite a plus de valeur qu'un simple exercice de calcul : c'est quand on examine le bien-fondé de certaines hypothèses, en comparant les résultats du calcul avec les données de l'expérience. S'il y a concordance, on en conclut que les hypothèses admises sont justes et qu'il n'entre en jeu aucun facteur dont on n'ait tenu compte dans le calcul ; s'il y a discordance, on peut conclure à la fausseté des hypothèses admises. Dans la plupart des cas, on pourra tirer la conclusion que l'indépendance présupposée des événements n'est pas réalisée effectivement ; souvent, la nature des écarts donne une indication sur la nature de la dépendance et permet de découvrir de nouvelles lois de la nature.

146. — La considération des probabilités étendue à l'astronomie peut servir et a servi à démêler des inégalités enveloppées dans les erreurs d'observation. Laplace par exemple appliqua ses méthodes au Système solaire (v. n° 11). L'un des phénomènes les plus remarquables du Système du monde est celui de tous les mouvements de rotation et de révolution des planètes et de leurs satellites dans le sens de la rotation du Soleil et à peu près dans le plan de son équateur. Un autre phénomène également remarquable est le peu d'excentricité des orbites planétaires et satellitaires, tandis que celles des comètes sont très allongées. Il y a, d'après Laplace ([82]), à parier plus de quatre milliards contre 1, que des phénomènes aussi remarquables ne sont point l'effet du hasard ; et

(82) Voir P.-S. Laplace, *Essai philosophique sur les probabilités.* Paris, 1814, p. 107-127 : « Du Calcul des probabilités, appliqué à la recherche des phénomènes et de leurs causes ».

ce grand mathématicien en a déduit une démonstration, devenue célèbre, qu'il doit y avoir eu une cause régulière à tout cela. Cette cause, Laplace la trouve dans la Nébuleuse primitive qui s'étendait primordialement au delà des orbes de toutes les planètes et qui s'est retirée successivement, par l'effet du refroidissement, jusqu'à ses limites actuelles.

J'ajoute qu'Antoine-Augustin Cournot (v. n° 20) a professé des idées analogues : les lois que nous découvrons dans la nature ne peuvent pas être dues au hasard, c'est-à-dire que les observations de coïncidences et de successions de phénomènes dont nous déduisons les lois naturelles ne peuvent pas être des coïncidences et des successions purement fortuites (83).

147. — Lorsqu'il s'agit d'ensembles infinis, il y a lieu d'introduire la notion de modèle ou d'échantillon d'un ensemble.

Définition 17. — Un ensemble fini S est dit un *échantillon* ou un *modèle* de l'ensemble M qui peut être fini ou infini, si S est constitué par un nombre suffisamment grand d'éléments *choisis au hasard* parmi les éléments de M. Il faudra dans chaque cas particulier examiner comment devra s'opérer ce choix fortuit pour que S soit un bon échantillon (v. n° 137, définition 15).

Le théorème de Jacques Bernoulli énonce une connexion entre la composition de l'échantillon S et celle de l'ensemble primordial M. En effet, supposons comme ci-dessus que l'ensemble M, fini ou infini, attribue au caractère B la probabilité mathématique p (v. n° 136). Supposons qu'après avoir extrait de M l'échantillon S, ce dernier ensemble attribue au même caractère B la probabilité mathématique p'. Le théorème de Jacques Bernoulli dit qu'un échantillon S dont le nombre d'éléments est suffisamment grand possède à peu près la même composition que l'ensemble primordial M dont il est extrait. Il s'ensuit que les probabilités mathématiques p et p' ne différeront que peu ou pas du tout l'une de l'autre. Le théorème de Jacques Bernoulli permet

(83) Voir G. Milhaud, *Le hasard chez Aristote et chez Cournot*, Revue de Métaphysique et de Morale, Paris, 1902, t. 10, p. 667-681.

même de préciser cette proposition et de calculer le nombre d'éléments que doit avoir l'échantillon S pour qu'on puisse s'attendre, avec une probabilité mathématique P fixée arbitrairement, à ce que la différence $p' - p$ ne dépasse pas certaines limites données d'avance. Cela précise un peu l'expression d' « échantillon suffisamment grand », qui figure dans la définition précédente.

Si un échantillon S n'a pas la composition à laquelle on serait en droit de s'attendre d'après des renseignements exacts sur l'ensemble primordial M, on en conclura que les éléments de S n'ont pas été choisis « au hasard » parmi ceux de M, dans le sens défini au numéro 137 ; donc, oubien on n'a pas scruté l'ensemble M que l'on croyait étudier, oubien il existait une règle ou une connexion qui influençait le choix, une « cause régulière » de la différence de composition entre l'ensemble et l'échantillon.

Exemple 1. — Avec 2 dés, on constitue théoriquement un ensemble infini M de coups possibles. Une série de 18 000 épreuves en est un modèle ou échantillon S. Si les diverses combinaisons de points de 2 dés se montraient dans S en proportions très différentes de celles qu'on est en droit d'attendre, par exemple si le sonnez y apparaissait avec une fréquence très différente de 1 : 36 (v. n° 4), mettons 5 000 ou 6 000 fois au lieu de 500, on pourrait en conclure à l'existence d'une « cause régulière » et dire que les dés sont pipés.

Exemple 2. — Prenons pour ensemble primordial M l'ensemble de tous les angles compris entre zéro et 45 degrés, ensemble infini qui a la puissance du continu linéaire. Comme échantillon S, prenons les angles de $0,1', 2', 3', 4', \ldots, 44°58', 44°59', 45°$. C'est un ensemble de $45.60 = 2\,700$ angles en progression arithmétique croissante dont la raison est 1 minute d'arc. Le critérium du choix est donc la grandeur des angles. Tout angle a une grandeur déterminée, mais le fait d'avoir justement une minute ou $2°36'$ ou n'importe quelle amplitude donnée d'avance constitue un caractère fortuit, n'étant pas compris dans la définition de la notion d'angle ni impliqué par elle. La table des 2 700 angles en question est donc bien un modèle de l'ensemble M, puisque c'est un ensemble fini et tiré « au hasard » de l'ensemble infini M.

148. — A propos du dernier exemple, Gauss ([84]) a fait un raisonnement tout-à-fait typique pour ce genre d'applications du calcul des probabilités, raisonnement qui montre que les méthodes stocastiques ne diffèrent pas, par leur nature, des méthodes utilisées dans d'autres branches des mathématiques appliquées, en mécanique ou en physico-chimie. Voici ce raisonnement :

a) On peut calculer *directement* log tg α par des séries, c'est-à-dire déterminer le logarithme de la tangente trigonométrique d'un angle donné α. D'autre part, comme tg $\alpha = \sin \alpha : \cos \alpha$ et par conséquent

$$\log \operatorname{tg} \alpha = \log \sin \alpha - \log \cos \alpha, \qquad \text{(VI 2)}$$

on peut calculer *indirectement* le logarithme de la tangente trigonométrique d'un angle α, en prenant la différence des logarithmes du sinus et du cosinus du même angle α.

b) Supposons tous les logarithmes calculés à n décimales près. L'erreur commise, ou l'écart, ne peut pas dépasser la moitié d'une unité de la dernière décimale conservée. L'intervalle $< -\frac{1}{2}, \frac{1}{2} >$ contient donc l'ensemble M de tous les écarts possibles. Cet ensemble M est continu et mesurable; sa mesure est $= 1$. Cherchons la probabilité mathématique d'un écart négatif (v. n° 136, définition 14). Le sous-ensemble T des écarts négatifs, c'est-à-dire des cas où le logarithme est évalué par défaut, a pour mesure $\frac{1}{2}$. Donc, la probabilité mathématique d'un écart négatif, dans log tg α abrégé à n décimales, est $= \frac{1}{2}$, en vertu de (VI 1). Il en est de même dans log sin α et dans log cos α abrégés à n décimales. Et la probabilité mathématique d'un écart positif, c'est-à-dire que le logarithme soit évalué par excès, est également $= \frac{1}{2}$ dans les mêmes conditions.

c) La probabilité mathématique que les erreurs soient de signes contraires, dans log sin α et dans log cos α abrégés tous deux à

(84) Carl Friedrich Gauss (1777-1855), *Einige Bemerkungen zu* Vegas *Thesaurus Logarithmorum*, Gauss' Werke, t. III, p. 257-264.

n décimales, est $= 1/2$, les quatre combinaisons possibles, $++$, $+-$, $-+$, $--$, ayant chacune la probabilité mathématique $1/2 . 1/2 = 1/4$ (v. les numéros 33 et 34).

d) La somme σ, obtenue par l'addition des valeurs absolues des erreurs commises en abrégeant log sin α et log cos α tous deux à n décimales, sera tantôt plus et tantôt moins de la moitié d'une unité de la dernière décimale conservée. Par un raisonnement analogue au précédent, on trouve que la probabilité mathématique pour que cette somme σ soit supérieure à $1/2$, est égale à $1/2$.

e) Dès lors, la probabilité mathématique pour que, log sin α et log cos α étant abrégés tous deux à n décimales, les deux erreurs ainsi commises soient à la fois de signes contraires et donnent, par leurs valeurs absolues, une somme σ supérieure à $1/2$ est, en vertu du théorème des probabilités composées, $1/2 . 1/2 = 1/4 = 25\ \%$.

Or, quand cet événement composé se produit, l'égalité (VI 2) ne peut plus avoir lieu *dans la dernière décimale conservée*, si log tg α a été calculé directement. On arrive donc à la conclusion suivante : *Si, dans une table de logarithmes à n décimales des fonctions trigonométriques, tous les logarithmes ont été calculés directement*, il faut s'attendre à ce que, dans 25 % des cas, la dernière décimale de log tg α ne soit pas égale à la différence des dernières décimales de log sin α et de log cos α ; en d'autres termes, *il faut s'attendre à ce que l'égalité* (VI 2) *soit trouvée fausse dans 25 % des cas.*

149. — Gauss fut le premier à faire ce raisonnement. Sitôt l'idée conçue, il la réalisa, car il était frappé de voir les marques du hasard dans les mathématiques elles-mêmes. Prenant les tables de Köhler où log tg α est calculé directement de minute en minute, Gauss se mit à comparer les dernières décimales, de façon à vérifier l'exactitude de l'égalité (VI 2), et cela pour les $60 . 15 = 900$ angles compris entre 30° et 44°59′. En dénombrant, Gauss trouva 225 fois, c'est-à-dire dans exactement 25 % de tous les cas, une différence dans le sens indiqué ci-dessus. C'est donc une concordance remarquable.

Si, en faisant un dénombrement analogue dans une autre table de logarithmes des fonctions trigonométriques, on trouve pour la

différence en question une fréquence relative très différente du 25 %, en particulier si une telle différence ne se produit jamais, on devra, comme Laplace dans sa cosmogonie (v. n° 146), conclure à l'existence d'une « cause régulière » qui entrave le hasard. Par exemple, au lieu des tables de Köhler utilisées par Gauss, qu'on prenne les tables de Véga, *Thesaurus logarithmorum completus*. On constatera que l'égalité (VI 2) est toujours satisfaite jusque dans les dernières décimales conservées ; l'existence de la « cause régulière » est donc certaine. Il est très naturel d'en tirer la conclusion que, dans cet ouvrage, les logarithmes des tangentes trigonométriques n'ont pas été calculés directement, mais qu'on les a déterminés par soustraction, conformément à (VI 2), et que seuls les log sin x et les log cos x ont fait l'objet de calculs directs.

L'expérience de Gauss prouve que les échantillons extraits des tables de logarithmes peuvent correspondre à un très haut degré aux théories stocastiques (v. n° 147, définition 17).

Cet exemple est particulièrement propre à montrer le rôle du calcul des probabilités dans l'investigation scientifique. Les méthodes stocastiques permettent, étant données certaines hypothèses, d'en tirer des conclusions concernant la fréquence de certains événements. Si l'observation confirme les résultats du calcul, on y verra une confirmation des hypothèses faites, et l'on pourra s'attendre à voir les événements se dérouler ultérieurement de la même façon. Si au contraire l'observation est en désaccord avec les résultats du calcul, on pourra seulement tirer la conclusion que les circonstances réelles diffèrent à certains égards de celles qui ont servi de point de départ à la théorie, autrement dit : qu'il y a une cause troublant l'effet du hasard. Les calculs ne sauraient conduire plus loin ; et la question : « De quelle nature peut bien être cette cause ? » n'est plus du ressort du mathématicien (comparer avec le n° 185).

150. — Il n'est pas dit que tous les événements qui paraissent dus au hasard se prêtent à des investigations stocastiques. Pour qu'on puisse appliquer les méthodes du calcul des probabilités à des phénomènes du monde réel, leurs fréquences relatives, quand le nombre des épreuves ou des observations est considérable, doi-

vent suivre la loi de Gauss, c'est-à-dire doivent donner lieu à des écarts qui se répartissent conformément au théorème de Jacques Bernoulli (v. chap. IV). Telle est la condition nécessaire pour que les calculs aient une valeur objective. Il est parfaitement indifférent qu'il y ait ignorance partielle ou connaissance complète de tous les détails, hasard subjectif ou non (v. n° 124, définition 11 et n° 143). *On devra considérer comme fortuits les événements dont les fréquences relatives obéissent au théorème de Jacques Bernoulli.*

On peut d'ailleurs s'en rendre compte à priori. La théorie des chances, en tant que science abstraite, opère avec la notion du hasard dans le sens logique (v. n° 135, définition 8) On arrive de là aux théorèmes fondamentaux de cette théorie par une suite de déductions mathématiques, en suivant une voie purement abstraite et sans avoir besoin à aucun moment de recourir à quelque représentation concrète de l'idée de hasard.

Il résulte de tout cela que *la probabilité mathématique est parfaitement compatible avec le savoir*, contrairement à ce que pensent beaucoup de savants. Il est faux de dire que l'ignorance au moins partielle est une condition nécessaire pour qu'on puisse appliquer les théorèmes du calcul des probabilités. Le hasard dans le sens logique subsisterait même pour une intelligence omnisciente capable d'assigner les causes physiques qui nous échappent et de prévoir les faits particuliers qui se dérobent à nos prévisions. Les caractères de la succession fortuite tiennent à l'absence de connexion logique. Je vais en donner trois exemples.

151. — *Comme premier exemple*, prenons les dénombrements faits par Gauss dans les tables de logarithmes (v. n° 149). C'est un très bon exemple à l'appui de notre thèse ; d'abord, parcequ'il ne peut s'agir ici que de hasard dans le sens logique ; ensuite, parce qu'on peut calculer la valeur de la probabilité mathématique sans qu'il y ait ignorance ni de conditions nomologiques ni de circonstances ontologiques (v. n° 117). La notion de « logarithme calculé directement, à n décimales près, des tangentes trigonométriques des angles de zéro à 90 degrés, de minute en minute, » délimite univoquement un ensemble fini de $60.90 = 5400$ nombres ; c'est l'ensemble M de la définition 14 posée au numéro 136. Tous ces

nombres sont bien déterminés pour une valeur donnée de n, par exemple $n = 7$, car il y a une relation parfaitement déterminée qui relie la septième décimale du logarithme d'un nombre à ce nombre lui-même. Dans ce choix de 5 400 nombres, rien n'est laissé au hasard, puisqu'une loi inflexible enchaîne tous les chiffres. Mais chacun de ces nombres, étant une « chose complète » dans le sens du numéro 130, est indéterminé par rapport à toutes les propriétés qui ne sont pas énoncées explicitement dans la définition. L'une de ces propriétés non énoncées est « la coïncidence de la dernière décimale avec celle du log tg calculé par l'égalité (VI 2) ». L'apparition de cette dernière propriété chez un des 5 400 éléments de l'ensemble M est donc un hasard dans le sens logique, bien qu'il ne puisse y avoir le moindre doute sur les éléments qui la possèdent et ceux qui ne la possèdent pas. Du moment où l'algèbre gouverne tout dans cet exemple, l'ignorance même partielle est exclue ; on peut calculer exactement, dans chacun des 5 400 cas particuliers, si la différence en question existe ou non. Malgré ce savoir complet, la probabilité mathématique est assignable : sa valeur est 1/4 ; elle est confirmée par l'expérience (v. les numéros 148 et 149). Cela prouve bien que le hasard logique subsiste malgré la connaissance la plus complète de tous les détails.

152. — *Comme second exemple*, envisageons la répartition des divers chiffres 0, 1, 2, 3, ..., 9, dans une table de logarithmes. Supposons qu'on prenne un très grand nombre de logarithmes consécutifs et que l'on considère leurs $n^{ièmes}$ décimales, par exemple la troisième décimale de chacun d'eux. On verra que les dix chiffres 0, 1, 2, 3, ..., 9 sont répartis à peu près également sur cette liste. Henri Poincaré dit : « Un instinct invincible porte à le penser, et d'ailleurs on peut le vérifier *a posteriori* » ([85]). On part donc de l'idée que l'apparition de tel ou tel chiffre à l'endroit d'une décimale donnée du logarithme, mettons à l'endroit de la troisième décimale, est un événement fortuit ; en d'autres termes : on peut tirer la parallèle avec le schéma de l'urne. D'une urne contenant dix billets marqués respectivement des numéros 0, 1, 2, 3, ..., 9,

(85) Voir p. 313 de l'ouvrage cité dans la note (81).

on en tire un au hasard ; après l'avoir inscrit comme troisième décimale du premier logarithme, on remet le billet dans l'urne, on brasse à nouveau son contenu et l'on répète l'opération. Le numéro sorti au second tirage fournit le chiffre pour la troisième décimale du second logarithme, celui sorti au troisième tirage fournit le chiffre pour la troisième décimale du troisième logarithme et ainsi de suite. Si l'on remplissait de cette façon successivement les troisièmes décimales des logarithmes des nombres envisagés, la probabilité mathématique de l'apparition de n'importe quel chiffre à la dite troisième décimale serait bien égale à 1/10 (v. n° 23, problème 3) ; de plus, en vertu de la loi des grands nombres, chacun des dix chiffres 0, 1, 2, 3, ..., 9, devrait apparaître, en moyenne, avec la même fréquence relative ; et ces nombres de fréquence devraient donner lieu à des écarts obéissant au théorème de Jacques Bernoulli (v. les numéros 80, 93, 96).

Joseph Bertrand a eu l'idée et la patience de vérifier le fait expérimentalement sur la septième décimale des 10 000 premiers logarithmes de la table de Véga citée au n° 149, table où les logarithmes sont indiqués avec 10 décimales. Si la série des 10 000 décimales en question présentait réellement les caractères d'une succession fortuite, il faudrait s'attendre, sur les 10 000 épreuves ainsi constituées, à trouver mille fois le chiffre 0, mille fois le chiffre 1, mille fois le chiffre 2 et ainsi des autres, ou peu s'en faudrait. Vérification faite, la septième décimale fut trouvée être 990 fois le chiffre zéro ; 997 fois le chiffre 1 ; 993 fois 2 ; 1 012 fois le chiffre 4 etc.

Il y a plus ; on voit apparaître les caractères de la loi de Jacques Bernoulli : Partagez les 10 000 nombres empruntés à la table de Véga en 10 séries de 1 000 nombres chacune ; vous aurez, pour la septième décimale du logarithme, 10 séries de 1 000 chiffres. Dans chacune de ces séries, la fréquence relative du chiffre zéro est voisine de la probabilité mathématique 1/10, mais il y a un écart. Les 10 séries donnent 10 écarts ; prenez-en la moyenne. Puis calculez les carrés des 10 écarts et prenez la moyenne de ces 10 carrés ; enfin, divisez la deuxième moyenne par le carré de la première ; vous devrez trouver approximativement le nombre $\frac{\pi}{2} = 1{,}570\,796\ldots$ (voir les formules du hasard, n° 96). Le dénombrement fait par

Bertrand a donné 1,561 pour le chiffre zéro. Le même calcul fait pour le chiffre 1 a donné 1,598. La moyenne des deux résultats est 1,579 ; les trois premiers chiffres sont exacts.

153. — *Comme troisième exemple*, citons une étude stocastique faite par Mr Henri Bruns. Dans les susdites tables de Véga, *Thesaurus logarithmorum completus*, où se trouvent calculés à dix décimales les logarithmes de tous les nombres de cinq chiffres, chaque page contient 5 colonnes à 60 logarithmes. Comptez dans chacune de ces colonnes le nombre des zéros finaux, c'est-à-dire des zéros à la dixième décimale (Au lieu du chiffre 0, on pourrait d'ailleurs prendre l'un quelconque des chiffres $d = 0, 1, 2, 3, \ldots, 9$, arbitrairement choisi, mais fixe ; Mr Bruns a choisi $d = 0$). Relevez, pour les 1 000 premières colonnes par exemple, combien d'entre elles ne comprennent aucun zéro final, combien comprennent 1 seul zéro final, combien en comprennent 2, combien 3, puis 4 etc., enfin 60. Vous trouverez une répartition en tous points semblable à celle que présentent des événements fortuits ; c'est comme si les zéros finaux avaient été distribués au hasard parmi les dixièmes décimales de ces logarithmes ; les écarts que présentent leurs fréquences relatives obéissent sensiblement au théorème de Jacques Bernoulli. Pour mettre cette répartition en parallèle avec le schéma de l'urne, imaginons une urne $U(1, 9 ; 10)$ [v. n° 33, définition 8], l'unique boule blanche représentant le zéro final et les 9 boules noires les autres chiffres de la dixième décimale. Comme chacune des 1 000 colonnes est à 60 lignes, 1 épreuve devra consister en une série de 60 tirages de cette urne, la boule tirée étant chaque fois remise dans l'urne et le contenu brassé à nouveau. L'événement E étant l'apparition du chiffre zéro, l'événement F l'apparition de l'un des 9 autres chiffres (v. n° 23 et n° 63), on a $p(E) \equiv p = 1/10$; $p(F) \equiv q = 9/10$. Le développement de $(p + q)^n = (0{,}1 + 0{,}9)^{60}$ permet de comparer les résultats du dénombrement avec ceux du calcul (86). Le nombre des épreuves

(86) V. p. 8 et p. 279-280 de l'ouvrage de Bruns cité dans la note (40) ; voir aussi le numéro 229, p. 423-425, de l'ouvrage de Czuber cité dans la note (46).

est $s = 1000$. En désignant par x le nombre des zéros finaux dans une colonne, on voit que x peut théoriquement prendre les valeurs $x = 0, 1, 2, 3, \ldots\ldots, 59, 60$, puisque chaque colonne comprend 60 nombres. Le nombre des colonnes avec x zéros finaux étant désigné par $k(x)$, le dénombrement effectif dans les susdites tables de Véga a fourni les nombres de la deuxième colonne du tableau ci-dessous (p. 216) ; ce sont les fréquences absolues k (v. n° 52). La colonne (3) donne les fréquences théoriquement attendues, $Fr(x)$, et la colonne (4) les fréquences relatives $k : s$ réellement constatées. On voit que les écarts entre l'observation et le calcul ne dépassent pas une dispersion normale.

On voit par les colonnes (3) et (5) de ce petit tableau que, dans les susdites tables de Véga, la fréquence relative des colonnes qui ont 7 zéros finaux est $= 0{,}134$; celle des colonnes qui ont au plus 7 zéros finaux est $= 0{,}747$, de sorte que la probabilité mathématique à postériori, c'est-à-dire la probabilité empirique, qu'une colonne prise au hasard contienne *exactement* 7 zéros parmi les 60 chiffres finaux, est $= 0{,}134$; et la probabilité empirique qu'elle contienne *au plus* 7 zéros est $= 0{,}747$ ou presque 3/4.

Il ressort encore de ces résultats que le nombre moyen des zéros finaux dans une colonne est 6,024 ; en effet, on compte le zéro final

$$1.6 + 2.36 + 3.78 + 4.149 + 5.161 + 6.183 + 7.134 + 8.114 + 9.74 + 10.34 + 11.19 + 12.10 + 13.0 + 14.2 = 6024$$

fois dans les 1000 premières colonnes.

La probabilité empirique que la 10ième décimale du logarithme d'un nombre entier choisi au hasard soit zéro devient dès lors $6{,}024 : 60 = 0{,}1004$ alors que la théorie des chances donne $0{,}1000$; et cela subsiste pour une décimale quelconque d'un rang suffisamment élevé.

Répartition des zéros finaux dans une table de logarithmes.

Nombre des zéros x	Fréquence absolue constatée $k(x)$	Fréquence absolue calculée Fr (x)	Fréquence relative $\frac{k}{s} \equiv p(x)$	Somme des probabilités $S_x \equiv \sum_{n=0}^{n=x} p_n$
(1)	(2)	(3)	(4)	(5)
0	0	1,8	0,0000	0,000
1	6	12,0	0,006	0,006
2	36	39,3	0,036	0,042
3	78	84,3	0,078	0,120
4	149	133,6	0,149	0,269
5	161	166,2	0,161	0,430
6	183	169,3	0,183	0,613
7	134	145,1	0,134	0,747
8	114	106,8	0,114	0,861
9	74	68,6	0,074	0,935
10	34	38,9	0,034	0,969
11	19	19,6	0,019	0,988
12	10	8,9	0,010	0,998
13	0	3,6	0,000	0,998
14	2	1,4	0,002	1,000
15 à 60	0	0,6	0,000	1,000
	1000	1000,0	1,000	

V. — *La probabilité mathématique et le principe des zones de comportement.*

154. — Nous avons montré ci-dessus que le principe des zones de comportement est à la base philosophique du théorème de Jacques Bernoulli et de la loi des grands nombres et qu'il permet de les expliquer admirablement l'un et l'autre (v. n° 128). Des suppositions qui ont, dans le sens du numéro 122, des zones comportaires indifférentes, primaires et égales, ont aussi des probabilités mathématiques égales. En vertu du même principe, une hypothèse dont la zone de comportement comprend une très grande partie du spaciom global possible *Sp* (v. n° 120) doit être considérée comme très probable, et une possibilité dont le domaine de comportement n'est qu'une partie infime du dit domaine *Sp* est très invraisemblable. Cela reste vrai même quand la probabilité étudiée n'a pas de valeur numérique assignable, ce qui arrive notamment quand les zones comportaires des diverses possibilités ne sont pas complètement indifférentes dans le sens du numéro 120, définition 6, ou qu'elles ne sont pas comparables entre elles ou, surtout, quand elles ne sont pas primaires.

On voit par là que la portée du principe des zones comportaires dépasse de beaucoup le domaine des probabilités numériques. La logique du bon sens applique ce principe journellement. Toutes les fois que la réalisation d'un événement « serait un grand hasard » et paraît très invraisemblable, on verrait, en scrutant la question, qu'il s'agit d'une conjecture basée en dernière analyse sur ce principe. Dans la théorie générale de la probabilité, le principe des zones comportaires ainsi étendu se place à côté de l'analogie et de l'induction. C'est une source logique particulière de probabilités, indépendante des autres. Rappelons que pour donner à ce principe toute la précision dont il est susceptible, il faut le compléter par la théorie des ensembles.

155. — On peut considérer d'un autre point de vue encore le principe des zones comportaires, en le mettant en parallèle avec le

principe de causalité. Lorsque nous conjecturons sur l'avenir en partant de ce que nous avons constaté dans le passé, ces expectatives se basent sur le principe de causalité, sur le postulat que toute l'évolution est dominée par des lois immuables et que les mêmes causes produisent toujours les mêmes effets. Mais quand nous avons remarqué certaines régularités dans la fréquence des événements et que nous nous attendons à les voir se reproduire avec « à peu près » la même fréquence relative, cette expectative se fonde sur le principe des zones de comportement et se justifie par lui. Malgré leur différence profonde, le principe de causalité et celui des spacioms comportaires ont ceci de commun qu'ils apparaissent comme des principes primordiaux, c'est-à-dire irréductibles à d'autres et pouvant servir de base à des prévisions de l'avenir.

L'explication déduite du principe des zones de comportement a un sens différent de celle déduite du principe de causalité ; elle n'en est pas moins une explication. En effet, le besoin intellectuel de comprendre est satisfait, quand l'esprit voit que l'arrivée d'un phénomène, étant données certaines circonstances générales, est liée à la réalisation non pas de tous les cas possibles impliqués par les dites circonstances générales, mais de l'immense majorité d'entre eux. La portée de ce principe s'étend bien au delà des probabilités mathématiques. Il englobe le vaste domaine des événements semblables. Considérons-en quelques

Exemples. — D'après les relevés annuels de la poste, un certain pour cent des lettres expédiées, pour cent à peu près constant d'une année à l'autre, portent une adresse insuffisante. On peut présumer que le même phénomène se reproduira l'année prochaine ; et pourtant, on ne saurait imaginer une loi naturelle qui amènerait inéluctablement ce fait. Que A, B et C aient inscrit des adresses incomplètes sur une partie de leur correspondance n'est vraiment pas une raison, ni pour que X, Y et Z en fassent de même l'année prochaine ni pour qu'ils ne le fassent pas. Il n'y a, entre les cas individuels, aucune dépendance, aucune connexion qui pourrait garantir une fréquence relative à peu près constante ; et l'on ne connaît rien actuellement qui entraînerait avec une nécessité absolue le résultat présumé pour l'année prochaine. Dès

lors, on se dit que certaines circonstances générales restent à peu près constantes et que le résultat le plus probable en est la répétition de l'oubli en question avec une fréquence relative déterminée.

Comment rendre compte de la certitude avec laquelle, appuyés sur des constatations passées, nous attendons la réalisation de certains phénomènes avec une régularité quasi mathématique, ne souffrant que de rares exceptions ? Par exemple : qu'une graine semée dans de bonnes conditions germera ; ou que la température moyenne de l'année prochaine, en un lieu déterminé, sera comprise entre 9^0 et 10^0 ; ou qu'en dénombrant les diverses lettres de l'alphabet sur une page des œuvres en prose de Gœthe, on trouvera sur 1000 lettres 160 fois la lettre *e* ? Et comment expliquer que nous nous attendons à des exceptions ou, du moins, que nous en admettons la possibilité ?

C'est ici qu'intervient le principe des zones comportaires. Au principe de causalité à expliquer les lois ; celui des zones de comportement rend compte des exceptions.

VI. — *Probabilité et régularités statistiques.*

156. — Le principe des zones de comportement a l'avantage d'écarter une fausse interprétation de la loi des grands nombres. Tant que l'on ne connaît pas d'autre principe d'expectative que le déterminisme absolu qui relie l'effet à sa cause, on est tenté d'interpréter exclusivement dans ce sens les régularités statistiques constatées dans les phénomènes sociaux, biologiques, démographiques, psychologiques, économiques etc. On arrive alors au résultat que le hasard *doit* se compenser et les diverses possibilités s'égaliser *à la longue* ; on arrive à conclure que les divers cas individuels, pourvu qu'on leur donne le temps de se réaliser, doivent nécessairement se produire avec une fréquence telle que la moyenne se conserve toujours.

Cette interprétation erronée est très répandue, surtout quand il s'agit des jeux de hasard. Déjà Laplace a consacré un chapitre spécial de son *Essai philosophique sur les probabilités* (v. n° 11) aux « Illusions dans l'estimation des probabilités ». Voici ce qu'il

écrit (p. 157 158 de la 2e édition, Paris, 1814) : « Lorsqu'à la loterie de France (87), un numéro n'est pas sorti depuis longtemps, la foule s'empresse de le couvrir de mises. Elle juge que le numéro resté longtemps sans sortir doit, au prochain tirage, sortir de préférence aux autres... » C'est une erreur ; la probabilité mathématique est toujours = 1/18, si le brassage se fait consciencieusement à nouveau après chaque tirage (v. n° 39).

Les joueurs ne se rendent pas compte de l'indépendance absolue des parties successives. Lorsque pile est amenée 9 fois de suite, ils se figurent qu'à l'épreuve suivante, face a plus de chances de se montrer. C'est une grande illusion ; en effet, oubien la pièce est bonne dans le sens du numéro 38, oubien elle ne l'est pas. Dans le premier cas, la probabilité mathématique est toujours = 1/2. Dans le second cas, mettons que le côté pile se trouve favorisé par suite de la constitution physique irrégulière de la pièce et qu'il ait été amené déjà 9 fois de suite, pour cette raison ; à l'épreuve suivante, c'est encore pile qui a le plus de chances d'arriver, et non le résultat opposé attendu par le joueur.

Lorsque rouge sort plusieurs fois de suite à la roulette, les joueurs considèrent l'arrivée de noir comme presque certaine à la partie suivante, parceque, disent-ils, le hasard, à la longue, *doit* compenser les écarts qu'il produit. Cette interprétation déterministe de la loi des grands nombres impliquerait que les parties successives fussent liées de telle façon que les résultats passés exerçassent une influence sur les résultats à venir, dans le sens d'une compensation des écarts. Dans cette hypothèse, où s'arrêter ? Les parties jouées la veille et l'avant-veille et depuis le début de la saison et l'année passée, et celles jouées depuis l'existence de la Maison et depuis des siècles, et dans la localité voisine et sur n'importe quel point du Globe, devraient exercer cette action étrange et mystérieuse sur les parties encore à jouer. Ce serait bien extraordinaire, incompréhensible et du point de vue mathé-

(87) Il s'agit de la Loterie génoise ou Lotto des nombres, venu d'Italie, introduit en 1757 à Paris. Voir Leonhardi Euleri Opera Omnia, éditées sous les auspices de la Société Helvétique des Sciences Naturelles, série I, t. 7, *Commentationes algebraicae ad theoriam combinationum et probabilitatum pertinentes*, ed. L. Gustave du Pasquier, 1923, préface, p. xxiii-xxvii.

matique tout-à-fait arbitraire. Au surplus, l'expérience prouve le contraire ; il n'y a pas d'origine des temps, il n'y a pas de liaison, les parties sont en fait absolument indépendantes et n'influent en rien les unes sur les autres.

L'illusion repose sur un sophisme : on allègue le théorème de Jacques Bernoulli ; mais il ne parle que de probabilités. Sur 20 000 épreuves, dit-on, à la roulette, la noire ne peut pas sortir plus de 12 000 fois, l'assertion de la Science est formelle et on ne l'a jamais constaté. Si donc les 10 000 premières parties avaient donné 6 000 noires, les 10 000 parties suivantes ont contracté une dette envers la rouge. « On fait trop d'honneur à la roulette », dit fort bien Joseph Bertrand, « elle n'a ni conscience ni mémoire ». Et il met à nu le sophisme en disant : « En supposant qu'à une rencontre inouïe succédera, pour la réparer, un nouvel écart de la règle, on n'efface pas l'invraisemblance, on la redouble ».

Laplace écrit aussi : « J'ai vu des hommes, désirant ardemment d'avoir un fils, n'apprendre qu'avec peine les naissances de garçons dans le mois où ils allaient devenir pères. S'imaginant que le rapport de ces naissances à celles des filles devait être le même à la fin de chaque mois, ils jugeaient que les garçons déjà nés rendaient plus probables les naissances prochaines des filles ». Laplace admet avec raison qu'ici encore, il y a illusion complète. D'ailleurs, où s'arrêter? Au lieu de se borner aux naissances dans Paris, il faudrait aussi considérer, si l'on adoptait ce point de vue erroné, celles qui se produisent dans la banlieue, dans tout le département, dans toute la France, dans toute l'Europe, celles de l'Asie, de l'Afrique, de l'Amérique et de la Terre entière. Comment admettre cette mystérieuse influence, que la naissance de quelque petit Chinois à Pékin puisse changer le sexe d'un fœtus à Paris ?

157. — Ces erreurs et ces illusions sont devenues moins négligeables lorsqu'il s'est agi des régularités dans le domaine de la statistique morale, de la criminalité etc. Des travaux de Quételet surtout ([88]), il semblait résulter que les actions d'un individu sont

([88]) Voir la note[100] du n° 185. Lambert-Adolphe-Jacques Quételet (1796-1874), mathématicien et statisticien belge, est le créateur de la statistique dite morale. Voir son ouvrage *Sur l'homme et le développement de ses facultés, ou Essai de physique sociale*, publié en 1835.

influencées et déterminées non seulement par ses circonstances individuelles et par ses motifs personnels, mais encore, et d'une façon tout-à-fait mystérieuse, par les actions de ses concitoyens connus ou inconnus et de tous les hommes en général. Il semblait que « le budget de l'échafaud et de la prison » exigeât chaque année son accomplissement rigoureux et que, lorsqu'il manquait par exemple quelques crimes au nombre moyen prédit par le calcul, certains individus, par une influence incompréhensible mais fatale, fussent incités à les commettre. On peut qualifier cette interprétation quételetienne d'erreur assez rapidement reconnue et dépassée ; mais il n'en reste pas moins vrai qu'il y avait là effectivement une lacune dans la compréhension. Même Mr Lexis, dans ses fameuses théories des événements sociaux, ne me paraît pas avoir comblé cette lacune, bien qu'il fasse ressortir avec netteté qu'il s'agit de cas individuels tout-à-fait indépendants les uns des autres. Il se borne à observer que de pareilles régularités peuvent se produire sans qu'il y ait dépendance des événements entre eux, et il le prouve par quelques exemples : les jeux de hasard, la fréquence relative d'une lettre déterminée de l'alphabet dans une langue, et d'autres encore.

Pour expliquer ces régularités statistiques, on peut en appeler au principe de causalité et au déterminisme rigoureux, aux lois de l'hérédité et de la physiologie, lois qui entraînent une certaine uniformité dans les générations successives, aux habitudes et aux besoins qui maintiennent nécessairement une certaine constance ; mais pour véritablement comprendre les exceptions parfois déconcertantes présentées en fait par ces régularités, le mieux est d'avoir recours au principe des zones de comportement ; voir le numéro 155. D'ailleurs, il serait vain de chercher à délimiter exactement ce qui doit être attribué au premier principe et ce qui est à mettre sur le compte du second.

Les choses se présentent autrement lorsque nous ne connaissons pas encore assez bien les circonstances générales ni surtout les zones de comportement correspondant à la réalisation de tel ou tel phénomène. Ainsi, on ne peut pas expliquer, avec la même netteté que dans le jeu de pile ou face (v. n° 38), comment il se fait que, dans un grand pays, *le taux annuel de nuptialité*, c'est-à-dire la fréquence relative des mariages au cours d'une année,

varie seulement entre des limites étroites. C'est encore la considération des zones comportaires qui apportera ici une sorte d'explication ; avec cette différence toutefois : au lieu de mettre les phénomènes à expliquer en rapport avec des faits sûrement connus par ailleurs, on les met en connexion avec des hypothèses plus ou moins vraisemblables (v. n° 160).

158. — Le principe des zones comportaires n'est pas de nature expérimentale ; c'est un principe apriorique au même titre que le principe de causalité. Ni l'un ni l'autre ne se démontrent. Il est vrai que la confiance avec laquelle nous l'appliquons croîtra, conformément aux lois psychologiques, au fur et à mesure que l'expérience confirmera plus souvent les présomptions basées sur ce principe. Mais ce n'est point là une démonstration logique, car le principe ne contient rien qui soit homogène à des lois empiriques ou qui puisse être déduit d'elles.

Si nos expectatives ou nos conjectures n'étaient pas confirmées par les faits, on en conclurait tout d'abord, comme on le fait pour le principe de causalité, que les hypothèses spéciales dont on était parti sont fausses ; dans l'espèce : que les rapports de grandeur des zones de comportement diffèrent de ceux qu'on avait admis, ou que des erreurs se sont glissées dans la détermination des probabilités. Et même si toutes les hypothèses étaient reconnues justes et que l'expérience ne confirmât pas la conjecture, on dirait qu'on a « rencontré un hasard très curieux » plutôt que d'abandonner le principe. Une exception, même fréquente, n'obligerait donc nullement à lui substituer autre chose.

159. — Nous pouvons, sinon imaginer, du moins concevoir et, par une sorte de passage à la limite, penser une intelligence si infiniment supérieure à la nôtre, si infiniment puissante et bien informée de la Nature, qu'elle connaîtrait en détail, à tout moment, la position et le mouvement de chaque atome de l'Univers. Le passé, le présent et l'avenir ne formeraient plus pour elle qu'un seul tout, un univers à quatre dimensions. La définition de la probabilité mathématique subsisterait pour cette intelligence surhumaine (v. n° 136), mais conjecture et présomption deviendraient sans objet et, pour cette intelligence omnisciente, le calcul des probabi-

lités n'aurait pas le même genre d'intérêt que pour nous. Or, un tel état de perfection dans la connaissance est inaccessible à nous autres mortels, à cause du nombre fantastiquement énorme des éléments : atomes, molécules, électrons etc., qui interviennent déjà dans le moindre phénomène physique. Un cube d'un micron de côté seulement contient quelque 34 milliards de molécules d'eau. Il en résulte que si l'on voulait calculer complètement le phénomène le plus simple, on se heurterait d'emblée à des difficultés insurmontables. Les chimistes qui ont essayé de faire de « l'astronomie moléculaire », comme par exemple Berthelet au début du XIXe siècle, n'ont pas tardé à reconnaître que le problème était insoluble à cause de sa complexité. Une comparaison va le rendre plausible (89). Mathématiciens et astronomes ne sont pas encore parvenus à résoudre d'une manière générale et complète le problème de trois corps en présence soumis à l'action de la gravitation newtonienne ; c'est donc encore par approximations successives seulement que les mouvements de notre système planétaire sont déterminés. Or, en physique, quand on a devant soi 1 cm^3 de gaz à zéro degré et à la pression ordinaire, on se trouve en présence de 3.10^{19}, soit 30 milliards de milliards de molécules. Va-t-il falloir, pour résoudre ce problème, poser 30 milliards de milliards d'équations différentielles simultanées ayant chacune des milliards de milliards de termes, représentant les actions mutuelles de toutes ces molécules les unes sur les autres ? « La seule pensée d'un tel problème ferait reculer d'effroi le mathématicien le plus résolu », dit Mr Ch.-E. Guye (89). Et en effet, si l'on voulait s'occuper de chacune de ces molécules pendant une seconde seulement, cela prendrait dix milliards de siècles, de sorte qu'il faudrait consacrer à cet examen environ 20 milliards de vies humaines. Cet exemple suffit à montrer l'impossibilité pratique absolue de résoudre de pareils problèmes d'une façon tout-à-fait générale. Seule, la stocastique permet actuellement de pousser les investigations plus avant, et comme elle est basée sur le principe des zones comportaires, on voit par là la très grande importance de ce principe (cf. n° 165).

(89) Voir le livre de Mr Charles-Eugène Guye, *L'évolution physico-chimique* (Bibliothèque de synthèse scientifique), Paris, 1922, notamment le chapitre sur « L'évolution des phénomènes physico-chimiques et le calcul des probabilités », p. 30.

CHAPITRE VII

L'APPLICATION DU CALCUL DES PROBABILITÉS A LA PHYSIQUE

I. — *L'introduction du calcul des probabilités dans la physico-chimie.*

160. — Dans la théorie des erreurs d'observation, on a depuis longtemps utilisé le calcul des probabilités pour *ajuster* les résultats bruts fournis par l'expérience (méthode des moindres carrés). Mais il n'y a pas si longtemps qu'on a eu recours aux probabilités mathématiques pour *expliquer* certains phénomènes physiques et chimiques. L'idéal d'une explication scientifique est la réduction à une nécessité causale, au déterminisme absolu postulé par la Science. Là où cet idéal est irréalisable, l'esprit, ne pouvant arriver à la Nécessité, se tourne vers la Possibilité et vers la Probabilité qui en découle. Possibilité devient un succédané pour certitude, et probabilité un succédané pour vérité, d'autant meilleur que le degré de probabilité est plus élevé. Quand on ignore absolument les conditions ontologiques (v. n° 117), on en est réduit à faire des hypothèses et à examiner si les conséquences qu'on en déduit par les voies logiques et mathématiques sont en accord avec l'expérience. Cet accord est, en définitive, le seul critérium qu'on puisse invoquer en faveur de l'hypothèse. Des faits et des suppositions, voilà les deux catégories d'éléments qui doivent servir de points de départ pour expliquer nos sensations, pour rendre les phénomènes rationnellement compréhensibles et pour en prévoir d'autres.

161. — L'introduction, en physique théorique, de la notion de probabilité comme principe d'explication est liée au nom de Boltzmann (90). Bien que le sens de la vue et celui du toucher puissent faire supposer parfois que la matière a une structure continue, l'évolution de la science conduit toujours plus à admettre que cette structure est en réalité discontinue, granulaire. L'hypothèse atomique, une des plus anciennes de la science, est depuis longtemps le fondement des lois de la chimie. Après Daniel Bernoulli qui a jeté les premières bases scientifiques de la théorie cinétique, après Clerk Maxwell qui a le premier tenté d'expliquer les propriétés d'une masse gazeuse en la considérant comme formée d'un nombre fantastiquement grand de molécules en mouvement incessant, après les travaux de Boltzmann, après l'étude des dissolutions, des radiations, du mouvement brownien, l'hypothèse atomique a pris le caractère d'une nécessité physique. « Les atomes ne sont plus une fiction commode », disait Henri Poincaré devant la Société de physique, « il nous semble pour ainsi dire que nous les voyons, depuis que nous savons les compter. »

Cette notion de discontinuité n'est pas restée confinée aux particules matérielles. Elle s'est étendue : au domaine de l'électricité d'abord et l'on a vu apparaître *l'électron*, grain ou atome d'électricité négative, dont la réalité est maintenant attestée par un grand nombre de faits (expériences de Millikan) ; puis au domaine du magnétisme, depuis que Mr Pierre Weiss a introduit *le magnéton* ou grain de magnétisme (91) ; puis au domaine de l'énergie rayonnante où elle a donné naissance à l'hypothèse des quanta due à Mr Max Planck. Suivant cette hypothèse, l'émission de l'énergie rayonnante ne s'effectue que par petits paquets dits *quanta* ou grains d'énergie. Et dans ces différents domaines, les méthodes stocastiques se sont montrées utiles, voire même indispensables.

(90) Ludwig Boltzmann (1844-1906), *Vorlesungen über Gastheorie*, Leipzig, 1896-1898. Traduction française par A. Gallotti et H. Bénard, *Leçons sur la théorie des gaz*, 2 vol., Paris, 1902 et 1905.

(91) Pierre Weiss, *Les moments magnétiques des atomes et le magnéton*, dans l'ouvrage : *Les idées modernes sur la constitution de la matière*, Paris, 1913.

II. — *Résumé de la théorie cinétique des gaz.*

162. — Dans la théorie cinétique des gaz, on considère les molécules comme des sphères minuscules ou boules parfaitement élastiques, identiques entre elles, dont le rayon est de l'ordre de grandeur du dix-millionième de millimètre, boules animées d'un mouvement incessant obéissant aux lois de la Dynamique, s'entrechoquant mutuellement et choquant les parois du récipient qui contient le gaz, analogues à un essaim d'abeilles voltigeant en tous sens, mais emprisonnées dans un espace clos. Une même molécule subit plusieurs milliards de chocs par seconde; chaque choc modifie la direction ou la vitesse de la molécule ou tous les deux à la fois. Il est dès lors impossible de connaître le nombre et les vitesses des molécules qui choquent à un instant précis t_0 une certaine portion de la paroi. Mais *l'effet global* de ces chocs est mesurable : c'est ce qu'on appelle *la pression* du gaz sur cette portion de paroi.

On conçoit dès lors immédiatement : 1° que la pression doive augmenter quand on comprime le gaz, puisque cela augmente le nombre des chocs par seconde ; 2° que la pression doive diminuer quand on laisse le gaz se dilater, puisque les molécules, ayant plus de place, tamponnent moins souvent la paroi ; 3° que la pression doive augmenter quand on échauffe la masse gazeuse et diminuer quand on la refroidit, puisque, d'après la théorie, élever la température d'un corps, c'est augmenter l'intensité de l'agitation de ses molécules, donc leur communiquer des vitesses plus grandes et par conséquent augmenter le nombre des chocs par seconde (Lois de Mariotte et de Gay-Lussac).

On « explique » de la même façon comment il se fait que la pression d'un gaz soit la même dans tous les sens. Le plafond d'une chambre remplie d'air supporte la même pression que le plancher, tandis que, si au lieu d'air on remplissait la chambre d'eau, il y aurait pression sur le plancher et non sur le plafond ; c'est que la pression de l'eau est due à son poids, celle du gaz au contraire est due au tamponnement moléculaire contre les parois.

Dans l'impossibilité d'appliquer les lois de la mécanique rationnelle à des phénomènes d'une aussi inextricable complication

(v. n° 159), les physiciens se sont adressés au Calcul des probabilités. « L'appui qu'ils ont trouvé dans ce précieux auxiliaire a été, il faut le reconnaître, de nature à les récompenser du crédit qu'ils lui avaient accordé à priori », écrit Mr Guye(*). Seules, en effet, les méthodes stocastiques permettent de poursuivre les investigations dans un domaine où la Mécanique rationnelle reste impuissante. Les conditions nomologiques sont celles de la mécanique classique, tandisque les circonstances ontologiques sont tout-à-fait inconnues (v. n° 117).

163. — Avec l'hypothèse moléculaire, la théorie cinétique des gaz aboutit au résultat suivant : *L'état le plus probable d'une masse gazeuse enfermée dans un récipient* A *est l'homogénéité*, c'est-à-dire l'état où les molécules sont distribuées uniformément dans tout le volume. L'expérience le confirme avec une très grande précision. Voici comment la théorie cinétique peut l'expliquer.

Imaginons une cloison partageant le volume de notre récipient A en deux parties, B_1 et B_2, exactement égales entre elles. Nous pouvons supposer à cette cloison virtuelle une infinité de formes et de positions, de sorte qu'une molécule arbitrairement choisie parmi les N molécules enfermées dans A pourra se trouver de l'un ou de l'autre côté de la cloison ; si nous conduisons la cloison au hasard, la seule condition à remplir étant $B_1 = B_2$, chaque molécule prise individuellement a autant de chances de se trouver dans B_1 que dans B_2 ; en d'autres termes, la probabilité mathématique p pour qu'une molécule déterminée de notre masse gazeuse se trouve dans B_1 est $p = \frac{1}{2}$, et la probabilité mathématique q pour qu'elle se trouve dans B_2 est également $= 1/2$. Il en résulte que *la question de l'état le plus probable de notre masse gazeuse peut se ramener au schéma de l'urne*. En effet, imaginons une urne $U(r, r ; 2r)$ avec les conditions du numéro 33, définition 8, le nombre $2r$ des boules ou grains contenus dans l'urne étant extrêmement grand, même par rapport au nombre N des molécules gazeuses. Supposons ensuite qu'on effectue une série de N tirages de cette urne,

(*) Ch.-Eug. Guye, *l. cit.* (89), p. 31.

sortant k_1 fois boule blanche et n_1 fois boule noire, $k_1 + n_1 = N$. Enfin, au résultat final de cette série d'opérations faisons correspondre une répartition déterminée des N molécules entre les deux compartiments B_1 et B_2, d'après la règle suivante : dans B_1 autant de molécules qu'il y a de boules blanches tirées, donc k_1 ; dans B_2 autant de molécules qu'il y a de boules noires tirées, donc n_1. L'apparition des boules blanches et des noires, donc aussi la distribution des molécules entre les deux compartiments B_1 et B_2, pourra être soumise au calcul des probabilités. Comme ici $p = q = \frac{1}{2}$, le résultat le plus probable des N tirages est $k_1 = n_1 = N : 2$, en vertu de théorèmes stocastiques connus (v. n° 70). A ce résultat correspond une répartition des molécules telle que le même nombre s'en trouve dans B_1 que dans B_2. Et comme ce résultat est indépendant de la forme et de la position de la cloison virtuelle qui sépare les deux compartiments, on voit que l'état le plus probable du gaz est celui où les molécules sont distribuées uniformément dans tout le récipient A, c'est-à-dire l'homogénéité ; q. e. d.

Remarque. — L'équilibre d'une masse gazeuse étant un équilibre statistique, il faut d'après la théorie s'attendre à des écarts dans la distribution des molécules. Physiquement, de tels écarts signifient des *fluctuations de la densité du gaz*. Celles dont les probabilités ne sont pas nulles sont si infimes qu'elles échappent totalement à l'expérience même la plus minutieuse. La loi des grands nombres dépasse de beaucoup la finesse de nos procédés expérimentaux d'investigation. Il n'en reste pas moins vrai que la possibilité d'une fluctuation même appréciable subsiste en théorie. A côté du principe de causalité, nous retrouvons le principe des zones de comportement, et cette circonstance ne peut manquer d'avoir une grande portée au point de vue philosophique. Ce phénomène des fluctuations vient apporter un tempérament au déterminisme inflexible qui régit l'univers matériel.

164. — Un problème statistique concret va nous permettre d'éclaircir cette question. Il est certain que la vie physiologique d'un individu se déroule suivant les circonstances et le milieu

ambiant où il se trouve placé et que chaque phase en est déterminée inéluctablement. Et pourtant, l'observateur le plus perspicace, en suivant de seconde en seconde les manifestations compliquées de vie individuelle, n'arriverait pas à en déduire la loi ou l'allure de la mortalité humaine. Pour y parvenir, il faut faire abstraction des enchaînements nécessaires de l'évolution individuelle, il faut envisager les décès comme des événements dus au hasard et travailler, sur l'ensemble ainsi obtenu, selon les méthodes stocastiques développées aux chapitres III et IV. Elles ne permettent de prévoir que *le résultat global* de tous les facteurs de mortalité ; ce sera la loi de mortalité cherchée. La précision obtenue sera d'autant plus grande que le nombre des événements étudiés sera plus considérable.

Appliquons les mêmes procédés à notre masse gazeuse enfermée dans le récipient *A* (v. n° 163). Les chocs des molécules entre elles et avec les parois du récipient, quoiqu'obéissant aux lois de la dynamique, sont inconnus dans le détail. On les envisage comme dus au hasard, faisant abstraction des enchaînements nécessaires de l'évolution individuelle de chaque molécule. Les méthodes stocastiques permettent alors de prévoir, et avec une très grande précision, le résultat global de tous les chocs qui se produisent pendant une seconde. Ce résultat global exprime une loi physique. Celle-ci est donc en définitive un résultat statistique. Il dépend des valeurs moyennes de certaines grandeurs, notamment du nombre moyen des chocs par seconde et de la vitesse moyenne d'une molécule. Voilà comment le point de vue statistique s'introduit dans la physico-chimie, entraînant l'emploi des conceptions du calcul des probabilités.

165. — Les méthodes stocastiques ont un inconvénient : elles n'aboutissent qu'à des résultats moyens. Notre connaissance du monde réel inextricablement compliqué est moins complète que si nous pouvions pénétrer le mécanisme intime des phénomènes ; au lieu de la connaissance complète et de la certitude, ce sont des conjectures, des expectatives, avec la possibilité théorique de fluctuations. Il faut même présumer, en vertu du théorème de Jacques Bernoulli et de la loi des grands nombres, que les deux comparti-

ments égaux B_1 et B_2 du récipient A du numéro 163 ne contiennent pas à chaque instant exactement le même nombre de molécules. Bien que ces fluctuations de densité restent tout-à-fait imperceptibles, en théorie et dans la pratique, il n'en est pas moins vrai que, du point de vue philosophique, les lois physiques ou chimiques ainsi trouvées ne sont que hautement probables et les résultats seulement très approchés.

Malgré cet inconvénient, la connaissance que l'on peut ainsi atteindre est plus satisfaisante qu'elle pourrait sembler à première vue. Cela tient à deux choses. 1° Les lois statistiques trouvées se rapportent aux circonstances ontologiques seulement ; les conditions nomologiques sont connues ou peuvent l'être (v. n° 117). Tel est le cas par exemple dans la théorie cinétique des gaz. Les lois du choc des molécules gazeuses soumises à l'agitation thermique sont parfaitement connues. En regard de cela, la question de savoir en détail comment les diverses molécules sont placées et quelles sont exactement leurs vitesses, à un moment précis t_o, peut paraître d'intérêt secondaire, vu la complexité inextricable du moindre phénomène. 2° Les fluctuations théoriquement possibles et qui se produisent sans doute effectivement autour de la moyenne sont absolument négligeables, parcequ'inaccessibles à la mesure expérimentale, vu la petitesse de l'espace ou du temps dans lequel il faudrait les observer pour les percevoir. Les expectatives basées sur le résultat le plus probable sont donc infiniment vraisemblables.

Quand les éléments qui interviennent sont aussi nombreux que les molécules d'un gaz, l'erreur possible est beaucoup trop faible pour qu'aucun moyen humain d'investigation puisse la déceler. Une probabilité mathématique de l'ordre de grandeur de 10^{-30} ou de 10^{-100} doit être considérée pratiquement comme nulle. « En la négligeant, on ne sacrifie en tout cas rien à la rigueur scientifique », écrit Mr Emile Borel (92), « car il n'est pas de connaissance humaine qui soit vraie avec cette précision. Tout se passe donc comme si la loi approchée était rigoureusement exacte ». Voir la *Remarque* du numéro 60.

(92) Emile Borel, *Le hasard* (Nouvelle collection scientifique), Paris, 1914, p. 160.

III. — *La répartition des vitesses moléculaires. La loi de Maxwell-Boltzmann.*

166. — Pour caractériser l'état le plus probable d'une masse gazeuse enfermée dans un récipient A, il ne suffit pas de décrire les positions des molécules à un instant précis t_0, ou du moins leurs positions les plus probables, il faut indiquer également leurs vitesses, en grandeur et en direction. Cette étude est un peu plus compliquée que celle de leur répartition. La pression exercée par un gaz sur la paroi est proportionnelle à la température absolue. On est conduit ainsi à admettre que la force vive des molécules est elle-même proportionnelle à la température. Connaissant le nombre et les dimensions des molécules ainsi que la pression, on peut en déduire, par le calcul, la force vive moyenne et par suite aussi la vitesse moyenne. Ainsi, les molécules d'azote de l'air atmosphérique ont aux températures usuelles une vitesse moyenne de 400 à 500 m par seconde (1 800 km à l'heure). Et la vitesse moyenne d'une molécule est, à la température ordinaire, de 425 m/sec pour l'oxygène et de 1 700 m/sec pour l'hydrogène.

Maxwell fut le premier à voir clairement que les lois du choc élastique ne permettent pas que toutes les molécules aient à chaque instant la même vitesse. Sa *loi de répartition des vitesses* est l'une des plus belles et des plus fécondes applications du calcul des probabilités à la physique. Même si l'on admettait que les molécules fussent animées initialement de vitesses égales en valeur absolue et de directions variées, cet état serait modifié dès que chaque molécule aurait subi un choc ; un second choc amènerait de nouvelles modifications, etc. ; comme chaque molécule subit plusieurs milliards de chocs par seconde, on voit qu'en très peu de temps, l'inégalité la plus complète se serait produite dans les valeurs des vitesses. Devant le chaos de tels mouvements, la Mécanique classique reste impuissante. Pour pénétrer plus avant, il faut procéder comme au numéro 164 ci-dessus : il faut faire abstraction du lien de causalité et admettre que la distribution des vitesses soit la même que si elle était produite par le hasard.

En partant de cette hypothèse, Boltzmann a démontré à nouveau la loi de répartition de Maxwell, par un raisonnement un peu long, mais impeccable. La première démonstration donnée par Maxwell n'était pas rigoureuse, parcequ'il y appliquait à tort le théorème des probabilités composées (v. n° 34). Nous allons énoncer la loi de répartition de Maxwell sous trois formes différentes dont la première fait connaître une probabilité mathématique et les deux autres un nombre probable de molécules.

Première forme. — La probabilité mathématique qu'une molécule gazeuse arbitrairement choisie ait une vitesse comprise dans l'intervalle très petit $< v, v + \Delta v >$ est donnée par l'expression

$$\frac{4\ v^2.\Delta v}{\sqrt{\pi}.u^3}.e^{-\frac{v^2}{u^2}} \qquad \text{(VII 1)}$$

où $e \equiv 2,7\ 1828\ 1828\ 4\ldots\ldots$ est la base des logarithmes népériens, $\pi \equiv 3,\ 141\ 592\ 65\ldots\ldots$ le rapport d'une circonférence à son diamètre et $u \equiv$ la vitesse la plus probable des molécules, en d'autres termes la vitesse de la grande majorité des molécules de la masse gazeuse envisagée.

Deuxième forme. — Le nombre des molécules dont la vitesse est comprise dans l'intervalle très petit $< v, v + \Delta v >$ est proportionnel à

$$v^2.\Delta v\ e^{-h^2.v^2} \qquad \text{(VII 2)}$$

où h représente une constante qui dépend du nombre total des molécules et de la force vive de la masse gazeuse étudiée, donc de son poids, de sa pression et de sa température.

Troisième forme. — Dans l'état stationnaire d'une masse gazeuse, le nombre des molécules pour lesquelles les composantes de la vitesse suivant trois axes cartésiens rectangulaires OX, OY, OZ, sont comprises respectivement dans les intervalles très petits

$$< x, x + \Delta x >,\quad < y, y + \Delta y >,\quad < z, z + \Delta z >,$$

est donné par une expression de la forme

$$a.e^{-b^2.(x^2+y^2+z^2)}.\Delta x.\Delta y.\Delta z \qquad \text{(VII 3)}$$

où a et b représentent deux constantes positives qui dépendent du nombre des molécules et de la force vive totale de la masse gazeuse envisagée.

Définition 1. — On appelle *répartition normale* un état où les vitesses des molécules sont réparties conformément à la loi de Maxwell résumée par les formules de ce numéro.

167. — *Conséquences.* — 1) D'après la loi de Maxwell, des vitesses même très grandes sont encore possibles théoriquement, du moins la formule permet-elle de leur assigner une probabilité mathématique, tout comme la loi de Gauss admet l'existence théorique d'erreurs d'observation très considérables. Mais en fait, seules des vitesses qui ne s'éloignent pas trop de la vitesse la plus probable, ni dans un sens ni dans l'autre, ont quelque chance physique de se réaliser. Déjà les molécules ayant une vitesse dix fois plus forte sont excessivement rares, puisque la probabilité mathématique correspondante est de l'ordre de grandeur de e^{-100}, soit $3,72.10^{-44}$; malgré le nombre énorme de molécules que contient une masse même petite de gaz, il faut s'attendre à ne rencontrer aucune vitesse très près de zéro ou très supérieure à la plus probable. Les chocs continuels, se succédant avec la rapidité de plusieurs milliards par seconde, rendent impossibles ces vitesses extranormales. D'ailleurs, la formule de Maxwell n'exprime sous une forme mathématique que des conjectures, elle ne permet de mesurer que des expectatives ; il n'est pas dit que toutes les vitesses auxquelles elle assigne une probabilité mathématique se réalisent effectivement.

2) On déduit de la formule (VII 3) que les vitesses les plus fréquentes ne sont pas celles voisines de zéro, mais celles voisines de $1/b$.

3) L'équilibre d'une masse gazeuse étant de nature statistique, c'est-à-dire un équilibre mobile, il faut, d'après la théorie développée aux chapitres III et IV, présumer des écarts d'avec la distribution normale. Or, un écart de ce genre signifie physiquement une *fluctuation de la température du gaz*, puisque par définition la température d'une masse gazeuse est proportionnelle à la force vive totale de ses molécules, donc proportionnelle au carré de leur

vitesse moyenne. Les molécules lentes sont les froides, celles animées de grandes vitesses sont les molécules chaudes. Une petite partie de la masse gazeuse aura donc, d'après la théorie, une température un peu différente de la température moyenne de la masse, et ces fluctuations-là sont inversement proportionnelles à la racine carrée du nombre total des molécules. On doit donc s'attendre à l'existence de molécules dont la température peut être, pendant un temps très court du moins, pendant quelques milliardièmes de seconde, 5 ou 6 fois plus élevée que la température moyenne de la masse. Il s'agit bien entendu de températures absolues : si la température moyenne est de 27° centigrades = 300° absolus, « 5 à 6 fois plus » signifie donc 1 500° à 1 800° absolus, c'est-à-dire 1 200° à 1 500° centigrades. On n'a pas manqué de faire état de la présence momentanée de ces molécules chaudes pour expliquer certains phénomènes de la physico-chimie.

Ici encore, le phénomène des fluctuations vient apporter un tempérament au déterminisme absolu. A côté du principe de causalité, on retrouve celui des zones de comportement ; cela ouvre des horizons insoupçonnés aux considérations philosophiques sur l'évolution de l'Univers.

IV. — *Les phénomènes réversibles et le principe de Carnot.*

168. — Boltzmann a déduit de son hypothèse non seulement la loi de Maxwell, mais deux autres conséquences que l'on peut énoncer sous la forme de *théorèmes* :

4) Dans l'état stationnaire d'une masse gazeuse, c'est la répartition normale des vitesses qui subsiste.

5) Toute répartition des vitesses différente de la répartition normale tend spontanément, par le seul effet des chocs moléculaires, vers la répartition normale et finit toujours par y arriver.

Des critiques furent bientôt soulevées, particulièrement contre le dernier théorème. La plus importante est l'objection de Loschmidt (1876) ; on peut la résumer comme suit. Le dernier théorème de Boltzmann affirme que les transformations subies par une masse

gazeuse constituent un phénomène irréversible : passage d'un état de répartition anormale à l'état de la répartition normale des vitesses, jamais le passage inverse. Or, toutes les formules de la dynamique sont réversibles, c'est-à-dire resteraient justes si toutes les vitesses, à un moment donné, étaient simultanément changées de sens, en conservant leurs valeurs absolues et leurs directions. Donc, objecte Loschmidt, la proposition de Boltzmann ne peut pas logiquement être une conséquence des principes de la dynamique, lesquels en apparence du moins interviennent seuls dans la démonstration.

Cette contradiction, c'est encore le principe des zones comportaires qui permet de la faire disparaître. Pour le montrer, commençons par rappeler quelques faits et définitions.

169. — *Définition 2.* — Un phénomène est dit *idéalement réversible*, si toutes les phases de ce phénomène peuvent se reproduire successivement et dans le sens inverse. Imaginons par exemple qu'on ait cinématographié le phénomène et qu'on fasse mouvoir la pellicule photographique en sens rétrograde, en commençant par les dernières épreuves obtenues. Ce film présente l'image exacte de ce que doit être un phénomène idéalement réversible.

Exemples. — D'après les lois de la mécanique, un pendule dont on négligerait tous les frottements, pour ne tenir compte que de la pesanteur, doit osciller indéfiniment. Dans ce cas, l'oscillation de droite à gauche et celle de gauche à droite constituent deux phénomènes idéalement réversibles. — Le mouvement des planètes autour du Soleil pourrait être idéalement réversible, si l'on ne faisait intervenir que les forces gravitationnelles. Il suffirait d'imaginer qu'à un instant donné, par suite de quelque cataclysme et d'un coup de baguette magique, toutes les vitesses changeassent de sens. Le système planétaire repasserait alors, avec des vitesses égales mais de sens opposés, par toutes les phases qui ont précédé le cataclysme. — Ou bien encore : un cristal de glace se trouve dans un bain à zéro degré ; si l'on refroidit le bain, on verra le cristal croître ; si inversement on réchauffe le bain liquide, le cristal se dissoudra et pourra reprendre sa dimension et sa forme

primitives. — De même une masse liquide qui se transforme en vapeur et cette vapeur de nouveau en liquide. Voilà des types bien caractérisés de phénomènes réversibles.

170. — A côté des phénomènes réversibles, il y a toute une catégorie de phénomènes que l'on appelle plus spécialement *irréversibles*. — *Exemples.* — Un corps tombe sous l'action de la pesanteur ; il est arrêté dans sa chute par un obstacle. L'expérience enseigne que la force vive qu'il a acquise en tombant se transforme en vibrations et en chaleur qui se dissipe peu à peu dans le milieu ambiant. On aura beau inverser les conditions de l'expérience et fournir au corps une quantité d'énergie calorique exactement égale à celle développée lors de l'arrêt de sa chute. Jamais on ne verra le corps se mettre à vibrer puis à remonter à son niveau primitif. Le phénomène est très franchement irréversible. — Il en est de même de presque tous les phénomènes biologiques. L'évolution d'un être vivant en particulier se déroule, dans son ensemble, toujours dans le même sens : de la naissance à la mort. Jamais on ne voit un mort ressusciter ni un adulte se transformer progressivement en enfant.

La thermodynamique classique énonce comme suit la condition de réversibilité idéale. *Théorème* : Pour que l'évolution d'un système soit rigoureusement réversible, il faut que chacun des états successifs par lesquels passe le système diffère infiniment peu d'un état d'équilibre. — Or, pour qu'il en soit ainsi, les transformations doivent être infiniment lentes ; en y regardant de près, on voit que de telles conditions ne peuvent être pratiquement obtenues. Expérimentalement, tous les phénomènes, même les plus réversibles en apparence, ont une part plus ou moins grande d'irréversibilité.

171. — Pour montrer que la réversibilité idéale est une limite qui n'est jamais atteinte en réalité, reprenons les exemples du n° 169, empruntés à l'ouvrage de Mr Guye ([93]). Dans le cas du pendule, on doit tenir compte des frottements. Quelque minimes que soient leurs grandeurs, ils sont absolument inévitables et agissent à chaque instant, si bien que deux oscillations successives

([93]) Ch.-Eug. Guye, *l. cit.* (89), p. 34.

n'ont plus rigoureusement la même amplitude et que pour une même position, les vitesses ne sont plus rigoureusement égales et de sens opposés. — Dans le cas des planètes gravitant autour du Soleil, il y a les actions exercées par ces astres sur les masses fluides qui entrent dans leur constitution ; elles peuvent produire, comme on sait, un effet d'amortissement, excessivement lent il est vrai, mais qui représente la part d'irréversibilité inhérente à tout phénomène. Pour la Terre par exemple, le travail des marées se fait aux dépens du mouvement de rotation diurne et produit un ralentissement de l'ordre de grandeur d'un centième de seconde par siècle. Or, tous ces phénomènes thermiques continueraient à se produire et toujours dans le même sens, soit dans le sens d'un dégagement de chaleur, même après que le cataclysme susmentionné (n° 169) aurait renversé le sens de toutes les vitesses. La marche du système planétaire dans ses orbites newtoniennes n'est donc pas idéalement réversible.

En résumé, ni les phénomènes de la mécanique rationnelle complétée par la thermodynamique, ni ceux de la physique ou de la chimie ou de la biologie n'échappent à cette règle : *il n'y a pas en réalité de phénomène idéalement réversible*. Elle est si générale qu'on en a fait un principe devenu fameux, qui semblait devoir dominer notre connaissance de l'Univers ; on l'appelle *le deuxième principe de la thermodynamique* ou *principe de* Carnot. On l'a énoncé sous diverses formes dont voici les trois principales.

Première forme. — Quand on met en contact deux corps de températures différentes et qu'on abandonne le système à lui-même, la chaleur passe toujours du corps le plus chaud vers le plus froid, jamais en sens inverse.

Deuxième forme. — Pour toute machine thermique transformant d'une manière continue de la chaleur en travail mécanique, il faut disposer d'au moins deux sources de chaleur à températures différentes.

A signaler que de ce principe découle la proposition suivante. La quantité d'énergie que l'on peut transformer en travail mécanique va sans cesse en diminuant. On exprime ce fait en disant que *l'énergie se dégrade*, c'est-à-dire se dissipe en chaleur perdue.

172. — Rappelons que CLAUSIUS a donné une autre définition du principe de Carnot en introduisant la notion *d'entropie*. C'est une fonction dont la valeur ne dépend que de l'état actuel du système. On en connaît la valeur seulement à une constante additive près et, dès lors, seules les variations de l'entropie sont accessibles aux déterminations expérimentales. On ne connaît pas plus la valeur de l'entropie qu'on ne connaît la valeur d'une vitesse ; il faut toujours indiquer ou sous-entendre un point de repère : vitesse par rapport à la Terre ou vitesse par rapport au Soleil etc. (94). En adoptant cette nouvelle notion encore un peu vague, on peut énoncer le deuxième principe de la thermodynamique sous une

Troisième forme. — Pour un système isolé, la valeur de l'entropie ne peut jamais diminuer. Au fur et à mesure que les phénomènes se déroulent dans ce système, elle reste invariable ou va en augmentant. Mathématiquement, c'est une fonction à croissance monotone.

Cette propriété de l'entropie de varier toujours dans le même sens justifie l'appellation de « principe d'évolution » qu'on a donnée au second principe de la thermodynamique.

CLAUSIUS a montré que, lorsqu'un système isolé passe d'un état à un autre, l'entropie ne peut croître que s'il se produit à l'intérieur du système des phénomènes irréversibles. L'accroissement de l'entropie est donc une sorte de mesure de l'irréversibilité. Si tous les phénomènes qui se déroulent à l'intérieur du système isolé étaient idéalement réversibles, son entropie demeurerait constante. Comme c'est là un cas limite jamais atteint dans la pratique (v. n° 171), le système isolé évoluera toujours de manière à augmenter son entropie.

Or, comme « système isolé », on peut prendre la totalité de l'Univers accessible aux moyens d'investigation de la science humaine ; et alors se dresse l'objection de Loschmidt exposée à la fin du numéro 168. Nous allons l'examiner à la lumière des méthodes stocastiques appliquées à la théorie cinétique des gaz. On verra que c'est encore le principe des zones comportaires qui permet de l'écarter.

(94) D'après la théorie de NERNST, pourtant, la valeur de l'entropie serait nulle au zéro absolu.

V. — *Le déterminisme statistique et les fluctuations.*

173. — Envisageons un récipient A partagé en deux parties, B et C, par une cloison étanche aux gaz. Supposons B absolument vide et que dans C se trouvent une demi-douzaine de molécules gazeuses seulement. D'après la théorie cinétique (v. n° 162), elles sont animées d'un mouvement incessant et viennent heurter les parois qui les réfléchissent à la façon de billes de billard d'un poids négligeable et d'une élasticité parfaite. Aussi longtemps que le nombre des molécules reste petit, on pourrait prévoir par le calcul les diverses possibilités de répartition de nos molécules entre B et C et même déterminer la durée de chaque répartition possible, en supposant connus un état initial de distribution des molécules et de leurs vitesses ainsi que la forme du récipient A et en imaginant enlevée la cloison qui sépare B et C. Mais dès que le nombre des molécules est grand, que les chocs se comptent par milliards en une seconde, la possibilité de suivre le phénomène en détail par l'analyse mathématique prend fin (v. n° 159). On est alors obligé de faire abstraction de la causalité et d'envisager tout le phénomène, quoique le sachant rigoureusement déterminé, comme régi par le hasard (v. n° 164).

Cela change du tout au tout la méthode d'investigation La grandeur des espaces B et C entre seule en considération ; toutes les autres circonstances ontologiques : positions initiales des molécules, directions et grandeurs des vitesses initiales, formes des parois etc., ne jouent plus de rôle essentiel et même n'interviennent plus du tout. Pour employer la méthode stocastique, posons

$$\alpha \equiv \frac{B}{C+B}, \qquad \beta \equiv \frac{C}{C+B}. \tag{VII 4}$$

On voit que α est la probabilité mathématique qu'une molécule du gaz se trouve dans B, et que β représente la probabilité mathématique pour une molécule de se trouver dans la partie C. Pour plus de simplicité, supposons que le gaz envisagé soit un gaz parfait et désignons par N le nombre total des molécules gazeuses

enfermées dans le récipient. On sait que le développement de $(\alpha + \beta)^N$ d'après la formule du binôme renseigne sur les probabilités mathématiques de toutes les distributions possibles, voir ci-dessus le chapitre III, spécialement le numéro 63.

L'expérience apprend que, si l'on enlève la cloison de séparation entre B et C, le gaz se répand dans tout l'espace $B + C = A$ et qu'au bout d'un temps relativement court déjà, les analyses les plus minutieuses ne permettent pas de déceler la moindre différence, quel que soit l'endroit du récipient A où l'on prélève l'échantillon. Il s'établit donc dans A un état stationnaire par rapport à la pression et à la température. Malgré le chaos du mouvement moléculaire régi par le hasard, il n'arrivera plus jamais que toutes les molécules du gaz se trouvent rassemblées dans C et aucune dans B; nous avons là le type d'un phénomène irréversible. Et pourtant, *du point de vue purement logique*, aucune raison ne s'oppose à ce que la masse gazeuse envisagée ne repasse, à un moment donné, par ses états antérieurs parcourus en sens inverse.

Comment expliquer cela par le calcul des probabilités ? Comment cette théorie, avec l'incertitude qui lui est inhérente à cause du principe des zones de comportement, peut-elle faire irruption dans un domaine où le déterminisme le plus absolu semble régner en maître incontesté ?

174. — Parmi les $N + 1$ termes du développement de $(\alpha + \beta)^N$, le terme maximum est celui où le rapport des exposants de α et de β est le même que le rapport $\alpha : \beta$; autrement dit, le résultat le plus probable est celui où $\alpha . N$ molécules se trouvent dans B et $\beta . N$ molécules dans C (v. n° 71). Nous avons démontré que plus l'exposant N est grand, plus aussi la somme des termes qui avoisinent le terme maximal, l'emporte sur la somme de tous les autres (v. n° 81, γ et n° 128, *b*). Appliqué à notre cas particulier où l'exposant N est de l'ordre de grandeur de 10^{20}, donc énorme, cela signifie : Les modalités qui correspondent à l'état le plus probable de notre masse gazeuse ou à un état qui n'en diffère pas beaucoup, en d'autres termes, les modalités comprenant exactement ou à peu près $\alpha . N$ molécules dans le compartiment B et $\beta . N$ molécules dans C, avec une densité partout uniforme ou presqu'uniforme et

une répartition des vitesses strictement normale ou à peu près conforme à la loi de Maxwell (v. n° 166), ces modalités-là remplissent à elles seules presque tout le spaciom global de comportement *Sp* (v. n° 120). Si donc, faisant abstraction de l'enchaînement causal, on supposait que les mouvements de chaque molécule se produisissent au hasard, il faudrait s'attendre, avec une probabilité mathématique frisant la certitude, à voir se réaliser une répartition caractérisée par l'homogénéité, c'est-à-dire par une densité uniforme et une répartition normale des vitesses moléculaires, ou du moins à une répartition qui ne différerait que très peu de celle-là ou qui n'en différerait sensiblement que pendant des intervalles de temps excessivement courts, tout-à-fait inappréciables à l'expérience.

Par exemple, α^N représente la probabilité mathématique que toutes les molécules se retrouvent dans la partie *B* du récipient, donc que le gaz revienne à son état primordial. Quelle que soit la valeur de la fraction α définie par (VII 4), quand l'exposant *N* est de l'ordre des centaines de trillions, α^N devient inconcevablement petit. Théoriquement, une telle probabilité mathématique, quoique dépassant en petitesse tout ce que l'imagination peut se figurer, n'est pas nulle (v. n° 30) ; et si l'on prodiguait indéfiniment le temps, cette répartition-là pourrait finir par se réaliser, de sorte que la réversibilité du phénomène n'est pas logiquement exclue. Le théorème de Jacques Bernoulli permettrait même de calculer combien il faudrait exécuter d'épreuves et de séries d'épreuves analogues à celles décrites au numéro 163, pour qu'on puisse parier 1 contre 1, et même 9 contre 1, de tirer *N* fois de suite boule blanche et par conséquent de voir toutes les molécules gazeuses revenir dans le compartiment *C*, du moins pour un instant imperceptiblement court. Mais la durée que le calcul indiquerait comme nécessaire à cette opération dépasserait à tel point toute conception humaine qu'il est plus simple de dire : « Cette configuration-là ne se réalisera jamais ». — Sur la différence entre la théorie classique et la nouvelle, voir le numéro 176.

175. — Le raisonnement précédent, de même que celui du numéro 163, subsiste *mutatis mutandis*, si le compartiment *B*, au

lieu d'être vide, contient lui-même un gaz parfait différent de celui qui remplit C et sans action chimique sur lui. C'est alors le phénomène de la diffusion de deux gaz l'un dans l'autre. Le dit raisonnement aboutit au résultat que *l'état le plus probable* est un mélange aussi intime que possible des deux gaz, avec une densité uniforme partout et une répartition des vitesses conforme à la loi de Maxwell (v. n° 166). L'expérience prouve que cet état le plus probable se réalise effectivement, et les analyses les plus minutieuses ne permettent pas de déceler la moindre différence dans la composition du mélange, quel que soit l'endroit du récipient A où l'on prélève l'échantillon. Comme cette diffusion n'aura provoqué ni dégagement ni absorption de chaleur, l'énergie du mélange sera exactement celle qu'auraient les deux gaz séparés. Sur ce dernier point, les théories modernes sont d'accord avec les anciennes, mais non sur la portée et la signification philosophiques de la loi expérimentale de la diffusion.

176. — Il y a une différence profonde entre les conceptions modernes et la thermodynamique classique. Pour celle-ci, un gaz *doit* se distribuer également dans tout le volume qu'il occupe, un autre état est exclu ; c'est le déterminisme absolu résultant du principe de causalité. Les conceptions modernes, au contraire, s'appuyant sur le principe des zones comportaires, envisagent l'équilibre théorique seulement comme l'état le plus probable et font intervenir des fluctuations, c'est-à-dire considèrent comme certaine l'existence des écarts d'avec la configuration la plus probable. Le point de vue moderne s'inspire de l'analogie avec le schéma de l'urne. Il est en effet excessivement invraisemblable que, si l'on fait 1 milliard de tirages d'une urne contenant des boules blanches et des noires en nombre égal, on obtienne exactement 500 millions de blanches et 500 millions de noires ; il faudra s'attendre au contraire, avec une quasi-certitude, à ce qu'il y en ait de l'une des deux espèces un peu plus et de l'autre un peu moins de 500 millions. Certains phénomènes semblent mettre en évidence ces écarts fortuits et très passagers au sein d'une masse gazeuse.

177. — Boltzmann est allé encore plus loin dans cette voie, en établissant une formule indiquant la relation numérique entre la

probabilité mathématique d'un état donné et son entropie S (v. n° 172). Boltzmann a démontré que l'accroissement de l'entropie S est proportionnel à la différence des logarithmes des probabilités des deux états successifs considérés ; en formule :

$$S_2 - S_1 = k.(lnp_2 - lnp_1) = k.ln\frac{p_2}{p_1} \qquad \text{(VII 5)}$$

où S_1, p_1, désignent l'entropie et la probabilité du premier état de notre masse gazeuse, et S_2, p_2, les mêmes quantités quand elle a passé du premier au second état, lnx signifiant le logarithme népérien de x. Comme $S_2 > S_1$, puisque l'entropie va toujours en augmentant, il en résulte $p_2 > p_1$, en paroles : le second état est plus probable que le premier. Les phénomènes qui se déroulent à l'intérieur d'un système isolé ont donc pour effet de faire évoluer ce système vers des états de plus en plus probables.

L'intervention de la fonction logarithmique dans ce domaine se comprend par le théorème des probabilités composées Soient S_1 et S_2 les entropies de deux systèmes isolés G_1, G_2, par exemple de deux masses gazeuses absolument indépendantes l'une de l'autre. Soit S l'entropie du système constitué par l'ensemble de G_1 et de G_2 envisagés comme formant un seul et même tout, un système G ; soient encore p_1, p_2, p, les probabilités respectives de l'état actuel des deux systèmes G_1, G_2 et de leur ensemble G. On sait que $p = p_1.p_2$, c'est-à-dire que la probabilité mathématique p pour que les deux systèmes en question se trouvent simultanément dans les états envisagés est le produit des deux probabilités (v. n° 33). On sait d'autre part que l'entropie S de l'ensemble G est égale à la somme $S_1 + S_2$. Il s'ensuit que l'entropie S est une certaine fonction de la probabilité, $S = f(p)$, fonction qui doit être telle que

$$f(p_1.p_2) = f(p_1) + f(p_2) \qquad \text{(VII 6)}$$

Cette relation fonctionnelle devant être satisfaite quelles que soient les valeurs de p_1 et de p_2, on en déduit par un raisonnement mathématique la forme de la fonction inconnue f et l'on trouve

$$f(x) = c.lnx + \text{const.}$$

d'où découle immédiatement la formule (VII 5). On voit par là que

les conceptions de Boltzmann, appuyées sur l'hypothèse de la structure granulaire de la matière, permettent de réfuter l'objection de Loschmidt et d'expliquer par des phénomènes idéalement réversibles l'irréversibilité de tous les processus du monde matériel (v. n° 169). C'est parceque l'explication stocastique ne cherche pas à déterminer l'allure des phénomènes moléculaires rigoureusement définie par les lois de la mécanique (principe de causalité), mais cherche plutôt à déterminer l'allure la plus probable parmi toutes les allures possibles (principe des zones de comportement).

La méthode stocastique, parceque basée sur ce dernier principe, laisse ainsi la porte ouverte à une certaine indétermination de l'avenir, du moins en théorie Les fluctuations possibles font que, au moment de son équilibre, le système matériel envisagé n'est pas dans un état invariable, mais oscille autour de cette position d'équilibre théorique, par suite des écarts dans les mouvements moléculaires Les lois de la physico-chimie, que l'on s'était accoutumé à considérer comme fatales et inéluctables, sont remplacées par des lois statistiques admettant des exceptions excessivement rares et, de plus, d'une durée imperceptiblement courte. Au déterminisme absolu de la physique et de la chimie tend ainsi à se substituer une sorte de déterminisme statistique, moins absolu ; et cette modification profonde dans la signification même des lois expérimentales ne manquera pas d'influer sur la pensée philosophique.

VI. — *Autres applications du calcul des probabilités à la physique.*

178. — La théorie des probabilités intervient dans l'explication d'un très grand nombre d'autres phénomènes physico-chimiques. Ces applications sont si nombreuses dans ce domaine qu'il faudrait tout un ouvrage pour les exposer avec détail. Je me bornerai à mentionner deux exemples où le principe des zones comportaires remplace avantageusement le principe de causalité.

En premier lieu, les phénomènes de la radioactivité. Les 34 milliards d'atomes de radium qui se désintègrent chaque seconde, parmi

les $4 . 10^{21}$ atomes que contient un grâmme de cette substance, donnent lieu à des fluctuations et à des écarts auxquels on peut appliquer avec succès les méthodes stocastiques (95).

En second lieu, le déplacement des particules dans le mouvement brownien (96). Le mouvement désordonné d'agitation que l'on observe chez les très petites particules en suspension dans un fluide est attribué au choc que ces particules reçoivent de la part des molécules du fluide par suite de l'agitation thermique. Or, ce qui caractérise particulièrement ce genre de mouvement, c'est son irrégularité absolue. Considérons une de ces particules en suspension et notons sa position au bout d'un temps t. Si toutes les actions s'étaient compensées rigoureusement, la particule se retrouverait à sa position initiale. Mais ce cas est relativement très rare, quand le nombre des impulsions successives est considérable. On peut alors définir une fluctuation moyenne r (v. n° 103 et n° 104) et l'on arrive à la loi du mouvement brownien découverte par Mr Albert Einstein

$$r^2 : r'^2 = t : t' \qquad \text{(VII 7)}$$

Il y aurait à citer encore, dans ce domaine, les travaux de Mr Langevin sur les gaz ionisés (97) et sur sa très ingénieuse théorie du paramagnétisme (98); puis les études d'un grand nombre de savants sur le rayonnement, sur la dispersion diffuse de la lumière par les molécules, sur les gaz ultra-raréfiés (99), sur la théorie des quanta. On a même appliqué les méthodes stocastiques à l'astronomie et créé la science de la statistique stellaire (Cf. n° 146). La théorie des erreurs d'observation est basée en entier sur le calcul des probabilités, et l'on sait combien elle est importante et combien brillamment elle a fait ses preuves.

(95) Ladislaus von Bortkiewicz, *Die radioaktive Strahlung als Gegenstand wahrscheinlichkeitstheoretischer Untersuchungen*, Berlin, 1913.

(96) Voir à ce sujet le livre de Mr Jean Perrin, *Les atomes*. Edition revue et augmentée. Paris, 1924.

(97) Paul Langevin, *Thèse*, Paris, 1902.

(98) Paul Langevin, *Sur la perméabilité magnétique des gaz*, Journal de physique, t. IV, Paris, 1905, p. 678; Annales de chimie et de physique, 8e série, t. IV, 1905, p. 70.

(99) Voir le volume *Les idées modernes sur la constitution de la matière*, Paris, 1913.

On voit par les considérations qui précédent combien l'idée de loi statistique tend à se généraliser. La cause en est l'impuissance de la Science actuelle à pénétrer la complication inextricable des phénomènes. Bien que les méthodes stocastiques pénètrent de plus en plus les sciences dites exactes, la physique, la chimie, les applications les plus importantes et les plus étendues semblent être réservées à ces méthodes dans le domaine des sciences biologiques, sociales et psychologiques, dans ce qu'on appelle maintenant *la biométrique* ([99a]).

On peut même dire que plus la science est complexe, plus aussi l'application du calcul des probabilités y devient féconde.

(99 *a*) Voir par exemple Raymond Pearl, *Biometry and Vital Statistics*, Johns Hopkins University, Baltimore, 1922.

CHAPITRE VIII

LA PROBABILITÉ MATHÉMATIQUE BASÉE SUR LA THÉORIE DES COLLECTIFS

I. — *Le collectif empirique. Définitions préliminaires.*

179. — Les divers essais d'ériger le calcul des probabilités sur une base logiquement inébranlable et en même temps satisfaisante aux points de vue mathématique et philosophique, essais exposés dans les pages précédentes, partent tous du concept de « cas également possibles », source principale des objections soulevées contre la théorie des chances et la validité de ses résultats. Une théorie toute récente évite ce point de départ critiquable et écarte du même coup toutes les difficultés qui en dérivent (v. n° 21). Cette nouvelle théorie, objet de ce chapitre VIII, est basée sur le concept de hasard dans le sens logique, introduit au numéro 135, et sur celui de collectif mathématique que je vais définir.

Commençons par un exemple souvent cité : On mesure la taille d'un très grand nombre m d'hommes adultes habitant le même pays, appartenant tous à la même race et dont l'âge soit, mettons, compris entre 30 et 35 ans. Si les mesures sont faites au centimètre près, on entendra par « taille de n centimètres » toutes celles comprises entre $\left(n - \frac{1}{2}\right)$ et $\left(n + \frac{1}{2}\right)$ cm. On pourra traduire le résultat global de toutes ces mesures par une table ou liste où sera inscrit : combien il s'en est trouvé dont la taille était de 140 cm, combien de 141 cm, combien de 142 cm, de 143 cm et ainsi de suite jusqu'à 210 cm, si l'on suppose que les tailles extrêmes constatées aient été de 140 et de 210 cm. On fait donc abstraction de tous les autres caractères indi-

viduels, physiques, intellectuels ou moraux ; on en retient un seul, dit *caractère distinctif*, que l'on suppose mesurable ou numériquement exprimable à l'aide d'une variable x.

Définition 1. — L'ensemble ainsi constitué est *l'ensemble collectif* ou simplement *le collectif* ou *le syllepte* à étudier. Chacun des individus considérés est dit un *membre* ou un *exemplaire* ou, d'un terme plus général, un *élément* du collectif ou du syllepte. Leur nombre m est dit l'*extension* et la variable x est l'*argument* du collectif ou du syllepte.

Définition 2. — La table ou liste résultant des mensurations, observations et dénombrements, ordonnée, comme on vient de le dire, suivant les valeurs croissantes de l'argument, est dite la *liste de répartition* de l'ensemble collectif envisagé ; nous dirons aussi *la loi de répartition* ou la *fonction de fréquence* ou plus brièvement la *diadose* du collectif, en convenant que dans cette locution, le terme de « loi » n'implique pas nécessairement l'existence d'une formule correspondante d'analyse mathématique.

Au lieu de la taille, on pourrait prendre comme caractère distinctif la grandeur du thorax ou la largeur de la tête ou le poids de l'individu ou la couleur des yeux ou n'importe quel caractère physique.

Définition 3. — Quand il faudra préciser, je dirai *syllepte* si les éléments de la collectivité envisagée sont en nombre illimité, par exemple quand il s'agira d'un ensemble infini de nombres ou de tirages d'une urne supposés répétés infiniment souvent. *Collectif* sera le terme général. En particulier, un *collectif empirique* sera un ensemble fini ; il aura des exemplaires concrets, matériels, et s'opposera par là au *syllepte* et au *collectif mathématique* qui est un cas particulier du syllepte. La différence entre ces deux derniers concepts sera exposée ci-dessous (v. n° 188, β).

180. — Voici quelques *exemples* de collectifs qui se prêtent aux investigations stocastiques. 1) L'exemple du numéro 179.

2) Le collectif est l'état de la température en un lieu déterminé.

Les exemplaires ou éléments en sont les jours de l'année, et le caractère distinctif : la température moyenne diurne, c'est-à-dire la moyenne arithmétique du maximum et du minimum observés. La liste de répartition ou fonction de fréquence ou diadose indiquera le nombre de jours où la température moyenne diurne fut de -10^0, le nombre de jours où elle fut de -9^0, de -8^0 etc., jusqu'à 25^0 par exemple. Comme « température moyenne de n^0 », on comptera les températures moyennes comprises entre $(n - {}^1/_2)$ et $(n + {}^1/_2)$ degrés.

3) Le collectif est l'ensemble des mensurations d'une grandeur physique, les exemplaires du collectif étant les résultats de ces mesures ; caractère distinctif : l'erreur d'observation. La liste de répartition ou fonction de fréquence indiquera combien de mesures étaient affectées d'une erreur de -12, combien d'une erreur de -11, de -10, de -9 etc., combien de ces mesures ont donné un résultat exact, combien étaient affectées d'une erreur de $+1$, combien d'une erreur de $+2$, de $+3$ etc., où le signe $-$ indique erreur par défaut, et le signe $+$ erreur par excès. On entendra par « erreur de n » toutes celles comprises entre $\left(n - \frac{1}{2}\right)$ et $\left(n + \frac{1}{2}\right)$.

4) Le collectif est formé des naissances qui se sont produites pendant un laps de temps donné dans un pays déterminé, les éléments étant des séries de s naissances chacune ; caractère distinctif : la fréquence relative des naissances masculines.

5) Le collectif est la prose d'un idiome donné, les éléments étant les pages d'un livre écrit dans cet idiome, et le caractère distinctif le nombre de fréquence d'une lettre déterminée de l'alphabet.

6) Le collectif est constitué par des tirages d'une urne contenant des boules blanches et des noires. Les éléments ou exemplaires sont les séries de s tirages chacune ; caractère distinctif : la fréquence de boule blanche dans chaque série. La diadose ou loi de répartition indiquera donc : le nombre des séries ne contenant aucune boule blanche ; le nombre de celles contenant 1 boule blanche (et $s - 1$ noires), le nombre de celles où entrent 2 boules blanches (et $s - 2$ noires), puis 3 blanches (et $s - 3$ noires) etc., enfin le nombre des séries composées exclusivement de boules blanches.

Définition 4. — On voit par les exemples ci-dessus que l'argument d'un collectif ou d'un syllepte, la variable x, oubien ne peut prendre que des valeurs discontinues, en nombre fini ou infini, oubien doit être au contraire envisagée comme une variable continue. On étend ces qualificatifs aux collectifs eux-mêmes et l'on dit, suivant les cas, que le collectif est *discontinu* (exemples 4 à 6 ci-dessus) ou qu'il est *continu* (exemples 1 à 3).

181. — On peut condenser les développements précédents comme suit.

Définition 5. — Un ensemble d'objets ou de choses que l'on peut subsumer sous une notion générale et ordonner par rapport à un caractère distinctif mesurable ou numériquement exprimable constitue un *collectif* ou *syllepte*. Il est caractérisé par sa *liste de répartition* ou *loi de répartition* ou *fonction de fréquence* ou *diadose*. L'étude de ces collectifs au point de vue du calcul des probabilités constitue *la stocastique*.

On conçoit la très grande importance de cette science, puisque les notions de syllepte et de collectif peuvent s'appliquer aux domaines les plus divers.

La liste de répartition d'un collectif est formée de deux colonnes. La première contient les différentes valeurs $x_1, x_2, \ldots\ldots, x_\lambda, \ldots\ldots$ de l'argument x, ordonnées par grandeur algébriquement croissante. La seconde contient les nombres de fréquence correspondants, $z_1, z_2, \ldots\ldots, z_\lambda, \ldots\ldots$ A côté de chaque valeur x_λ se trouve écrit le nombre z_λ indiquant combien de fois la valeur x_λ du caractère distinctif se rencontre parmi les exemplaires du collectif étudié. La somme de tous les nombres z est évidemment égale à m, extension du collectif,

$$\sum_\lambda z_\lambda = m. \qquad \text{(VIII 1)}$$

On peut ajouter à la liste de répartition, sans en modifier en rien la signification, autant de x_λ que l'on veut, quand bien même les valeurs en question de l'argument ne figurent pas dans le collectif étudié, pourvu qu'on marque par zéro le nombre correspondant z_λ. On dit que ces x-là sont *vides* (Cf. n° 134, définition 5).

En fait, du moins lorsqu'il s'agit de collectifs continus, des limites bien déterminées ne sont pas assignables à l'argument x. Les valeurs extrêmes dépendent en effet du nombre m des observations effectuées. Par exemple, l'intervalle entre les tailles extrêmes est plus grand si l'on mesure 10 millions d'hommes que si l'on n'en mesure que dix mille. En ajoutant en pensée tous les arguments vides, on peut donc supposer que l'argument x parcourt tous les nombres de $-\infty$ à $+\infty$ et qu'à côté de chaque x_λ se trouve marqué le nombre correspondant $z_\lambda \geqq 0$. C'est ce qu'on suppose notamment quand on emploie les expressions de *loi de répartition* ou de *fonction de fréquence* ou de *diadose*.

On peut ajouter à la liste de répartition, à côté de la colonne z indiquant les nombres de fréquence des x_λ, une troisième colonne, celle des fréquences relatives y_λ, en posant (v. n° 52)

$$y_\lambda \equiv z_\lambda : m \qquad \text{(VIII 2)}$$

L'égalité (VIII 1) donne alors cette propriété de la diadose :

$$\sum_{\lambda=0}^{\infty} y_\lambda = 1. \qquad \text{(VIII 3)}$$

182. — Pour représenter graphiquement la liste de répartition ou la fonction de fréquence, on porte en abscisse les valeurs de x et en ordonnée les z correspondants. Ainsi dans le premier exemple (v. n° 179), les abscisses $x = 140, 141, 142, \ldots\ldots, 210$, expriment la taille, et les ordonnées correspondantes, z, marquent le nombre des individus possédant une taille comprise entre $(x - {}^1/_2)$ et $(x + {}^1/_2)$ centimètres.

Définition 6. — En joignant les extrémités de ces ordonnées par des segments de droite, on obtient une ligne brisée dite *la courbe caractéristique du collectif étudié.*

Son allure générale est la suivante : A gauche, venant de $-\infty$, elle se confond avec l'axe des X ; puis elle s'élève jusqu'à un maximum, abstraction faite d'un ou de plusieurs sauts éventuels ; puis elle s'abaisse de la même manière, pour finalement se confondre de nouveau avec l'axe des X jusqu'à $+\infty$.

Si l'on envisage cette courbe dans son ensemble, on voit que, malgré les sauts irréguliers marqués au coin du hasard, son allure traduit nettement une règle générale, une loi. De même que la surface terrestre, malgré la présence de montagnes et de vallées et de toutes sortes d'irrégularités grandes et petites, traduit dans son ensemble une surface sphérique ou ellipsoïdale. Pour une étude générale et systématique des ensembles collectifs, notamment quand il s'agit de sylleptes continus et homogènes, on est ainsi amené à *ajuster* les irrégularités fortuites, c'est-à-dire à aplanir les écarts dus au hasard. A la courbe caractéristique brute fournie par l'observation directe, on substitue une courbe ajustée. D'autres fois même, on se donne à priori la loi de fréquence sous une forme mathématique. Dans tous ces cas, on parle encore de *courbe caractéristique* du collectif.

Définition 7. — Dans le cas le plus simple possible, la courbe caractéristique a un seul maximum a partir duquel elle décroît d'abord lentement, puis très rapidement. C'est *la courbe de Gauss* ou *courbe des probabilités*, introduite au numéro 100. Son équation en coordonnées cartésiennes rectangulaires est

$$z = c.\ \mathrm{e}^{-k^2 (x - x_0)^2} \qquad \text{(VIII 4)}$$

où c et k représentent des constantes et x_0 la valeur de x qui correspond au maximum (fig. 2, p. 141). On dit dans ce cas qu'il y a *répartition normale* (cf. n° 166), ou encore : que « la répartition suit la loi exponentielle simple ou *loi de Gauss* ».

On rencontre cette allure dans d'innombrables cas, non seulement à propos des caractères somatiques de l'homme, de ceux d'animaux et de plantes, mais aussi quand les ensembles collectifs sont constitués par les objets ou les phénomènes les plus divers, même lorsqu'il s'agit de produits de la fantaisie artistique, comme l'a constaté Fechner en mesurant les dimensions des tableaux d'un grand musée de peinture.

183. — Quand la courbe caractéristique du collectif présente une asymétrie, notamment lorsqu'elle a deux maxima, il est souvent possible de la décomposer en une somme de deux courbes en

cloche. On peut alors conclure que le collectif étudié n'est pas homogène ou qu'il est constitué par un mélange de deux espèces dont chacune suit la loi de répartition normale. Voir la figure 3.

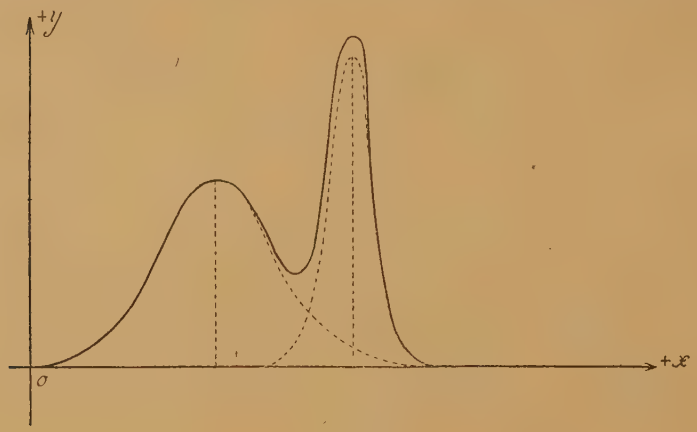


Fig. 3.

Exemples. — Si l'on mesure la taille de tous les habitants adultes d'une contrée où cohabitent deux races différentes, mettons des Blancs et des Japonais, et que l'on note les mesures sans distinguer les races, on constatera que le groupement autour de la moyenne ne suit pas la loi de Gauss. Il y a superposition de deux phénomènes distincts dont chacun suit la loi de répartition normale, mais dont la somme ne suit pas cette loi. Les tailles des Blancs se groupent autour de leur propre moyenne suivant la loi de Gauss, et de même les tailles des Japonais autour de leur moyenne ; on comprend que la moyenne générale n'ait plus aucune signification anthropologique. Ce fait biologique de la présence de deux races distinctes, s'il n'avait pas été connu, aurait pu être décelé par l'étude du collectif.

C'est ce qui est arrivé bien des fois. Ainsi, W. F. R. Weldon avait mesuré, pour en faire le rapport, la longueur et la largeur des exemplaires d'un lot très nombreux de crabes et avait tracé la courbe caractéristique du collectif, correspondant à ce caractère somatique. Il avait eu à sa disposition deux échantillons de crabes, les uns provenant de Plymouth et les autres de Naples. La

courbe fournie par l'étude des crabes de Plymouth était symétrique, celle des crabes de Naples ne l'était pas. En analysant cette dernière, Pearson trouva qu'on pouvait la représenter par la superposition de deux courbes de Gauss. On en peut conclure que les crabes recueillis à Naples étaient de deux types ou de deux variétés différentes, tandis que les crabes recueillis à Plymouth appartenaient tous à la même variété.

184. — Toutes les fois que le stocasticien aura des raisons de croire que les matériaux biologiques qu'il étudie sont un mélange de deux espèces différentes, l'une de petite taille, l'autre de plus grande taille, il cherchera à décomposer la courbe caractéristique du collectif envisagé en deux courbes plus simples. S'il y réussit, comme dans le cas de la *figure 3*, il sera tout naturellement porté à considérer la courbe ponctuée de gauche comme représentant la loi de fréquence de la taille pour la première espèce, et la courbe ponctuée de droite comme concernant la seconde. Il y a plus : La surface comprise entre chacune de ces courbes et l'axe des X donnera le nombre des individus de chaque espèce et par conséquent la proportion du mélange.

Le stocasticien pourra ensuite étudier d'autres caractères somatiques et examiner si ses conclusions sont confirmées ou non.

S'il a des raisons de penser que ses matériaux biologiques sont un mélange d'un plus grand nombre d'espèces ou de variétés différentes, il cherchera encore à décomposer la courbe du collectif étudié en courbes plus simples, c'est-à-dire à la considérer comme résultant de la superposition de 3 ou de 4 ou de 5 courbes de Gauss Et s'il y réussit, on pourra en conclure que l'ensemble étudié, quoique très homogène, est véritablement un mélange d'espèces différentes.

Ce sera d'autant plus probable que le mathématicien, de son côté, démontre qu'on ne peut pas décomposer n'importe quelle courbe plane c en une somme de courbes de Gauss, tandis qu'on peut décomposer n'importe quelle courbe plane c, pourvu qu'elle satisfasse aux conditions très générales connues sous le nom de conditions de Dirichlet, en une somme de sinusoïdes. Dès lors, quand la courbe caractéristique d'un collectif donné était

décomposable en un petit nombre de courbes normales de Gauss, on devait en conclure que cette possibilité de décomposition correspondait à quelque chose d'objectif. et cela avec d'autant plus de raison que le nombre des courbes normales composantes était plus restreint.

185. — Quételet, qui fut le premier à appliquer la loi de Gauss en dehors de la théorie des erreurs d'observation, croyait à l'universalité de la loi exponentielle introduite au numéro 182, définition 7. Quételet considérait cette loi non seulement comme un fait expérimental, mais comme une norme à laquelle l'expérience doit nécessairement se plier. C'est aller trop loin. Ayant prouvé la répartition normale en appliquant sa formule aux résultats des mensurations de thorax faites sur des soldats écossais, Quételet prit comme nouveau champ d'études les observations faites sur la taille de 100 000 conscrits français. Les données statistiques s'adaptaient assez bien aux résultats du calcul. Cependant, vu les normes légales, la théorie prédisait que 26 345 recrues devaient être réformées pour cause de taille trop petite. Or, en réalité, on en réforma 28 620, soit 2 275 de plus que le calcul n'indiquait. Que conclure de là ? Le bon sens dirait : « Cela prouve que pour les petites valeurs de l'argument, la concordance entre la théorie et l'observation n'est pas aussi bonne que pour les grandes ». Mais Quételet, attribuant à la répartition normale une sorte de réalité métaphysique, raisonne tout autrement ; selon lui, la discordance entre le calcul et l'expérience provient de ce que les dispositions légales ne furent pas toujours scrupuleusement exécutées. Quételet trouve donc que, dans 2 275 cas exactement, c'est par suite de tricheries que le conscrit ne fut pas reconnu apte au service militaire ([100]).

Il est assez curieux que cette théorie de Quételet proclamant la valeur universelle de la loi de Gauss, malgré l'abondance des

(100) Les idées de Quételet, malgré quelques erreurs et exagérations, contiennent en germe la biométrique moderne. C'est Quételet aussi qui a introduit la notion de *l'homme moyen*, notion critiquable, mais féconde. Il lui correspond certainement une réalité, un caractère de la race envisagée. Voir la note 88 du numéro 157.

matériaux qui auraient pu prouver le contraire, se soit ancrée si profondément chez les théoriciens. Bravais fut peut-être le seul, à ce moment-là, qui prit de suite le point de vue juste : à tout ensemble collectif, disait-il, correspond une loi de répartition particulière et, *a priori*, ce n'est pas nécessairement la loi exponentielle simple traduite par la formule (VIII 4). L'idée de Bravais, quelque juste qu'elle soit, passa inaperçue et il était réservé à Francis Galton et à Karl Pearson de fonder une nouvelle science, *la biométrique*, qui est une systématisation des recherches où l'on utilise les mesures faites sur des êtres vivants. La notion de collectif dans le sens du numéro 181 y joue naturellement un rôle capital.

186. — En prenant l'expérience seule pour base, on constate que l'interprétation d'une courbe caractéristique résultant d'observations statistiques peut devenir extrêmement difficile. En particulier le matériel biologique, surtout s'il est emprunté directement à la nature, peut fort bien ne pas contenir exclusivement deux espèces, mais encore, s'il s'agit d'espèces voisines, les résultats de leur

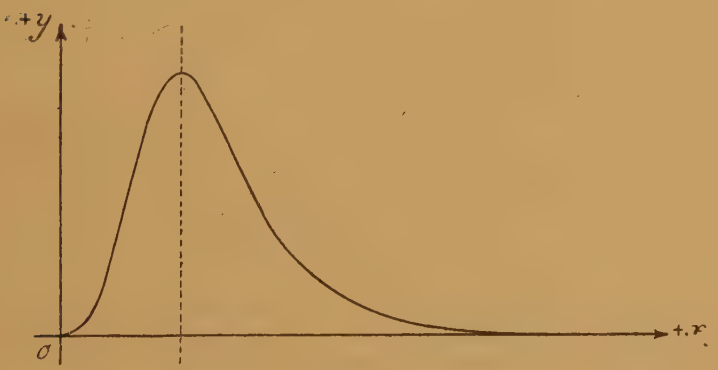


Fig. 4.

croisement ; et ces produits du croisement pourront suivre ou ne pas suivre la loi de Mendel relativement au caractère étudié. De plus, il arrive qu'une collectivité biologique se comporte par rapport à un premier caractère distinctif comme un ensemble homogène, mais

cesse de l'être par rapport à un autre caractère. On peut donc obtenir, par l'étude d'un premier caractère, une courbe en cloche bien caractérisée et conclure de là à l'existence d'une race pure. Puis, passant à l'étude stocastique d'un deuxième caractère sur le même ensemble biologique, obtenir une courbe à double sommet, ce qui ferait présumer l'existence de deux espèces différentes. C'est que dans des espèces voisines, certains caractères peuvent être communs et ces caractères, considérés isolément, peuvent faire croire à une espèce unique, alors qu'en réalité il s'agit tout de même d'un mélange de plusieurs espèces.

Enfin, un dernier point est à considérer. Même en opérant sur des lignées pures et sur des matériaux biologiques que l'on sait être parfaitement homogènes, il n'est pas dit que la liste de répartition du collectif, représentée graphiquement, donne après ajustement une courbe de Gauss normale, d'équation (VIII 4). *A priori*, on ne peut alléguer aucune raison péremptoire pour que la répartition des fréquences se fasse nécessairement, et quel que soit le caractère étudié, suivant la loi exponentielle simple. Du point de vue purement logique, on ne voit pas pourquoi l'augmentation d'un caractère devrait toujours avoir la même probabilité mathématique que sa diminution. Déjà Fechner ([39]), avec une grande persévérance, a accumulé des preuves démontrant l'insuffisance de la loi exponentielle simple pour la description de tous les collectifs empiriquement constatés. Et de toutes les recherches subséquentes, il résulte que la fonction de fréquence d'un collectif ou d'un syllepte peut ne présenter aucune symétrie et la courbe caractéristique correspondante être analogue à celle de la figure 4.

L'hypothèse d'une répartition normale a été corroborée, il est vrai, par l'étude de presque tous les caractères somatiques de l'homme. Mais ce n'est pas une preuve qu'elle soit applicable dans d'autres domaines, par exemple à des caractères intellectuels. D'un autre côté, lorsqu'on ne sait encore absolument rien de la fonction de fréquence d'un collectif, il est permis, et pour certains problèmes il est même nécessaire, de faire une hypothèse, ne fût-elle que provisoire. Arrivé aux derniers confins de la connaissance, le savant est libre de suivre n'importe quelle voie qui paraît devoir le conduire plus loin. Dans ce cas, on ne voit pas quelle

objection pourrait être soulevée valablement contre l'hypothèse d'une répartition normale, à condition de ne jamais oublier que les conclusions tirées d'une telle prémisse n'ont elles-mêmes qu'un caractère provisoire.

Résumons en disant : *a priori*, la fonction de fréquence d'un ensemble collectif ou d'un syllepte peut être quelconque, à condition de s'annuler pour les valeurs infiniment grandes de l'argument, de ne pas prendre de valeurs négatives et d'être univoque.

II. — *Le collectif mathématique. Nouvelle définition de la probabilité.*

187. — Je vais maintenant esquisser la nouvelle manière d'ériger le calcul des probabilités en discipline mathématique. On part de l'idée que la théorie des probabilités doit être à la fois une science de la nature et une discipline mathématique, au même titre que par exemple la géométrie ou la mécanique rationnelle. Elle doit donc se proposer comme but de rendre compte des relations entre les phénomènes observables, non pas en prenant tels quels les phénomènes du monde extérieur, mais en les idéalisant par abstraction. Comme toute théorie qui veut s'élever au rang d'une science, le calcul des probabilités doit partir de concepts logiques univoquement déterminés ; et pour poser ces notions primordiales, il doit tenir compte des choses qu'il s'agit de représenter. Mais une fois établi le système d'axiomes et de postulats auxquels ces concepts sont soumis, la Théorie doit procéder par voie purement déductive et arriver par des raisonnements impeccables aux théorèmes qui lui sont propres. Quant à savoir si la théorie est pratiquement utilisable ou non, seule l'expérience peut le décider.

Mr de Mises a érigé le calcul des probabilités sur un concept nouveau, logiquement bien défini, que l'on peut considérer comme idéalisation et formulation mathématique précise de la notion de « collectif empirique » employée dans la théorie stocastique de Fechner-Lipps ([39]) et de Bruns ([40]) (v. n° 179, n° 180). J'appellerai cette nouvelle notion un *collectif mathématique*. C'est la généralisation au cas où $m \to \infty$, des ensembles statistiques finis de m

éléments. On fait rentrer de cette façon dans la définition de la probabilité mathématique les propriétés infinitaires, c'est-à-dire précisément celles dont le collectif empirique est dépourvu. Ces propriétés infinitaires se manifestent : 1° dans la définition même de la probabilité mathématique comme limite d'une suite de rapports ou de quotients ; 2° par l'absence de régularité, par la fortuité avec laquelle le caractère distinctif appartient, ou non, aux éléments de l'ensemble considéré. En faisant de cette notion de collectif mathématique le pivot de la théorie, on tient compte d'une manière heureuse des besoins de la stocastique et de la biométrique.

188. — α) Désignons par (e) une suite infinie de choses idéales que, pour abréger, nous appellerons *éléments* et désignerons par

$$e_1,\ e_2,\ e_3,\ldots\ldots\ .,\ e_\lambda,\ldots\ldots\qquad (e)$$

Considérons en outre k variables réelles, $x_1, x_2, \ldots\ldots, x_k$, dont chacune soit susceptible de prendre une infinité de valeurs numériques.

Définition 8. — En langage géométrique, tout système de k valeurs particulières attribuées aux variables $x_1, x_2, \ldots\ldots, x_k$ est dit *un point* et les k valeurs particulières en question en sont *les coordonnées*. L'ensemble de tous ces points constitue un espace à k dimensions, un sp_k, dit *spaciom diacritique*.

Définition 9. — Faisons correspondre à chacun des éléments e_λ un système bien déterminé de valeurs particulières des variables $x_1, \ldots\ldots, x_k$. Nous dirons que ce système de valeurs particulières est *le caractère distinctif* de l'élément e_λ. En langage géométrique, nous faisons ainsi correspondre à chaque élément e_λ un point bien déterminé du spaciom diacritique sp_k.

β) Nous supposons que les éléments e_λ n'ont pas tous le même caractère distinctif, ni tous les éléments abstraction faite d'un nombre fini d'entre eux. A l'infinité des éléments correspond donc un ensemble infini de points différents. On voit que la suite (e), associée ou rattachée comme il vient d'être expliqué à l'ensemble des

caractères distinctifs, constitue l'analogue de la liste de répartition ou diadose du collectif empirique (v. n° 179). Du moment où il y a un nombre illimité d'éléments, nous avons un syllepte.

Mais celui-ci ne devient un collectif mathématique, C, que si la correspondance entre les éléments e_λ et leurs caractères distinctifs satisfait à deux postulats que je vais développer : celui des valeurs limites et celui de la fortuité.

189. — *Postulat I. — L'existence des valeurs limites.* — Soit A une portion quelconque du spaciom diacritique total sp_k (v. n° 188). Envisageons les n premiers éléments de la suite (e) et désignons par n_A le nombre de ceux d'entre eux dont le caractère distinctif fait partie de A. Le nombre n_A est donc la fréquence absolue et le rapport $n_A : n$ la fréquence relative du caractère distinctif en question, parmi les n premiers éléments de la suite (e) (v. n° 52). L'axiome I postule que la fréquence relative $n_A : n$, quand le nombre n des éléments considérés augmente indéfiniment, tende vers une limite finie et bien déterminée, non négative, qui peut dépendre de A et que je désigne par p_A ; en d'autres termes, l'axiome I postule l'existence de

$$\lim_{n \to \infty} \left(\frac{n_A}{n} \right) \equiv p_A \qquad \text{(VIII 5)}$$

et cela quelle que soit la partie A que l'on considère dans le spaciom diacritique sp_k.

Définition 10. — Cette limite, à condition que le postulat II ci-dessous soit également satisfait, est dite *la probabilité mathématique du caractère distinctif de A dans le collectif C.*

La probabilité mathématique n'a donc de sens précis que par rapport à un collectif bien déterminé C (comparer ceci avec le numéro 136). On dira, pour abréger, *la probabilité mathématique de A*, quand il n'y aura aucun doute possible sur le collectif C auquel cette probabilité se rapporte.

Définition 11. — *L'ensemble caractéristique M* est l'ensemble de tous les points qui peuvent constituer un caractère distinctif d'un élément ; en d'autres termes, cet ensemble M est formé,

parmi les points du spaciom diacritique total sp_k, par tous ceux que l'on peut faire correspondre à au moins un élément de la suite (e).

Dans le langage ordinaire, M est l'ensemble des « événements » ou encore l'ensemble des « cas possibles ».

Cet ensemble caractéristique M n'est jamais vide. Il peut comprendre la totalité du spaciom diacritique sp_k ou n'en comprendre qu'une partie ; il peut aussi se réduire à deux points, mais non à un seul, vu l'hypothèse du n° 188 β). De plus, à cause de la même hypothèse, il existe toujours au moins deux parties non vides de M à chacune desquelles correspondent une infinité de points de sp_k, et, parmi ces parties, au moins une pour laquelle la probabilité p_A n'est pas nulle. Je vais illustrer par un exemple les définitions précédentes.

Exemple. — Envisageons un corps ayant plus ou moins exactement la forme d'un cube dont les faces sont numérotées de 1 à 6. L'ayant secoué dans le cornet, on le lance au hasard ; le « point amené » est le numéro que montre la face supérieure. Considérons 1° comme « élément » de la suite (e) chacun des jets possibles de ce dé plus ou moins irrégulier et inhomogène ; 2° comme « caractère distinctif » de cet élément, le point amené. Nous avons donc une seule variable, x_1, de sorte que $k = 1$ et le spaciom diacritique sp_k est unidimensionnel. L'ensemble caractéristique M se compose de six points, dont les coordonnées sont 1, 2, 3, 4, 5, 6. La probabilité mathématique d'amener le point 6 par exemple est par définition

$$p(6) \equiv \lim_{n \to \infty} (n_6 : n) \qquad \text{(VIII 6)}$$

Cette limite existe par hypothèse ; et on peut la déterminer expérimentalement, comme on peut déterminer une constante physique quelconque, en faisant un nombre suffisamment grand d'observations. De cette manière se trouve définie la probabilité mathématique d'amener un point déterminé, même quand le dé utilisé est pipé ou irrégulier.

Remarque 1. — Si l'on admet la définition basée sur la théorie des ensembles (v. n° 136), c'est seulement à l'aide d'un théorème

qu'on arrive à la signification de $p(6)$ ou de $p(\lambda)$ par voie logique. On démontre en effet que *les probabilités mathématiques de l'apparition des diverses faces sont proportionnelles aux angles solides définis à l'aide de ces faces, en prenant pour sommet commun le centre de gravité du dé utilisé* (101).

Si l'on voulait se baser sur la définition classique, elle serait inapplicable au cas du dé non parfaitement régulier et homogène (v. n° 23, définition 1, *Remarque*, et n° 27). Il est en effet évident que « l'égale possibilité » des six faces du dé n'est pas autre chose qu'un postulat par lequel on définit un dé idéal. Pour savoir si, dans un cas concret, cette hypothèse ou ce postulat sont réalisés, on ne peut pas faire autrement que d'effectuer un très grand nombre N de coups et d'étudier les quotients $N_A : N$, car même si l'homogénéité de la matière utilisée pour la fabrication semble garantie, on reste dans l'incertain sur les influences que la plus petite asymétrie de la forme peut exercer *à la longue* sur les résultats des coups, si la différence de la forme réelle du dé d'avec la forme idéale du cube mathématique est tellement minime qu'elle reste imperceptible.

Dès lors, la définition (VIII 6) est la plus naturelle ; on a recours à elle pour vérifier pratiquement la régularité du dé.

C'est ce que fit entre autres l'astronome suisse Rodolphe Wolf (103). Ayant exécuté, avec un dé blanc et un dé rouge lancés simultanément, une première série de 20 000 coups, puis d'autres séries très nombreuses, en notant chaque fois le résultat, il a trouvé pour l'apparition des diverses faces les probabilités mathématiques suivantes, où $p(\lambda) \equiv$ probabilité mathématique d'amener le point λ ($\lambda = 1, 2, 3, 4, 5, 6$). (Voir le tableau page 264).

Remarque 2. — Il va de soi que l'égalité (VIII 3), valable pour les collectifs empiriques, reste vraie dans le cas des collectifs mathématiques. Si l'on partage le spaciom diacritique total sp_k d'une manière quelconque en ensembles partiels et que A parcoure tous ces ensembles partiels, la somme des probabilités mathématiques p_A ainsi obtenues est $= 1$.

(101) Voir F. M. URBAN, *l. cit.* (77), p. 91-92.

	dé blanc	*dé rouge*
$p(1) =$	0,16 230	0,17 035
$p(2) =$	0,17 245	0,18 155
$p(3) =$	0,14 485	0,15 880
$p(4) =$	0,14 205	0,14 580
$p(5) =$	0,18 175	0,17 240
$p(6) =$	0,19 660	0,17 110
	1,00 000	1,00 000

190. — Des définitions ci-dessus posées, on peut déduire de nombreux corollaires auxquels nous ajouterons quelques explications.

Théorème 1. — La probabilité mathématique est un nombre réel de l'intervalle fermé $< 0 ; 1 >$.

Théorème 2. — Dans un collectif mathématique C, la probabilité est une fonction définie pour tous les ensembles ponctuels du spaciom diacritique sp_k. Elle jouit de la propriété additive, c'est-à-dire qu'on a *le théorème d'addition*

$$p_A + p_B = p_{A+B} \qquad \text{(VIII 7)}$$

si les sous-ensembles A et B de l'ensemble diacritique n'ont aucun point commun et que $A + B$ représente l'ensemble somme de A et de B.

Définition 12. — Si l'ensemble caractéristique, M, se réduit à deux points, a et b, on dit que le collectif mathématique est *une alternative simple* ; $p(a)$ est dit la probabilité mathématique « de l'arrivée de a », et $p(b)$ est dit la probabilité mathématique « de la non-réalisation de a ». Dans ce cas, il suit de la définition que

$$p(a) + p(b) = 1 \qquad \text{(VIII 8)}$$

(v. n° 23, définition 2). On voit qu'on peut très bien subsumer sous les nouvelles définitions les locutions classiques usuelles dans le calcul des probabilités. D'une manière plus générale, on a le

Théorème 3. — Pour tout ensemble A contenant tous les points de l'ensemble caractéristique M, on a $p_A = 1$; pour tout ensemble A ne contenant aucun point de M, on a $p_A = 0$.

Pour définir *la diadose*, répartition des caractères distinctifs sur les éléments d'un collectif mathématique C, il est naturel de s'appuyer sur l'analogie avec les collectifs empiriques (v. n° 179). Posons donc comme

Définition 13. — On appelle *fonction de fréquence* ou *loi de répartition* ou *diadose* du syllepte C l'ensemble de toutes les valeurs p_A relatives à ce syllepte, c'est-à-dire l'ensemble des valeurs que peut prendre la probabilité mathématique p_A, quand A parcourt tous les sous-ensembles possibles du spaciom diacritique total sp_k. D'ailleurs, on peut se borner à envisager des ensembles partiels faisant partie de l'ensemble caractéristique M, puisque les autres points du sp_k ne fournissent aucune contribution aux valeurs de p_A (v. n° 189, définition 11).

Remarque 1. — Il est essentiel d'observer que la réciproque du théorème 3 n'est pas vraie ; en d'autres termes, $p_A = 0$ n'entraîne pas toujours l'impossibilité logique ni physique de l'événement étudié ; de même, $p_A = 1$ n'implique pas nécessairement la certitude. La probabilité mathématique peut en effet être nulle, en vertu de la formule (VIII 5), sans que l'ensemble des cas favorables ne soit vide, lorsqu'il s'agit des probabilités dites continues. Par exemple, dans la théorie gaussienne des erreurs d'observation, la probabilité mathématique que l'erreur commise soit comprise dans un intervalle dx est proportionnelle à la grandeur de cet intervalle ; dès lors, la probabilité mathématique que l'erreur soit rigoureusement égale à un nombre donné d'avance, α, est nulle. Et malgré cela, aux points de vue logique et physique, une erreur de grandeur α exactement est possible. Il y a là une différence de principe très marquée entre les théories classique et moderne (v. n° 30 et n° 174).

Remarque 2. — On peut se faire une image intuitive de la fonction de fréquence, en se rappelant que la somme de tous les p_A possibles est égale à l'unité (v. n° 189, remarque 2). Il suffit par exemple d'imaginer une répartition de la masse 1 sur les points de l'ensemble caractéristique M, répartition continue ou discontinue suivant la structure de M. Cette masse 1 ainsi répartie matérialise en quelque sorte la diadose ou fonction de fréquence.

191. — *Postulat II.* — *La fortuité du caractère distinctif.* — Soient A et B deux ensembles partiels du spaciom diacritique total sp_k, n'ayant aucun point commun ; supposons que les probabilités mathématiques p_A et p_B, déterminées conformément au n° 189, ne soient pas nulles. Dans la suite (e) des éléments e_λ (v. n° 188), commençons par biffer ceux dont le caractère distinctif ne fait partie ni de A, ni de B ; puis numérotons les éléments qui restent, en leur attribuant comme indices les nombres naturels 1, 2, 3, ..., n, ... De la suite illimitée ainsi obtenue, déduisons une nouvelle suite infinie

$$e'_1, e'_2, e'_3, \ldots\ldots, e'_\lambda, \ldots\ldots \qquad (e')$$

en disposant des indices « sans tenir compte des caractères distinctifs », opération que j'appellerai *un choix admissible*. Enfin, supposons déterminées, conformément au numéro 189, les probabilités mathématiques p'_A et p'_B. L'axiome II postule non seulement que ces limites p'_A et p'_B existent, mais en plus qu'elles satisfassent à la relation

$$p'_A : p'_B = p_A : p_B \qquad \text{(VIII 9)}$$

Définition 14. — *Opérer un choix admissible*, c'est former des suites partielles sans tenir compte des caractères distinctifs des éléments choisis ; cela revient à constituer de nouveaux collectifs mathématiques en partant d'un syllepte donné C.

Voici, à titre d'exemples, comment on peut procéder. — 1) On s'en tient à une règle arithmétique : ne prendre que les éléments d'indice pair ; ou ne prendre que ceux dont les indices appartiennent à une même progression arithmétique ; ou prendre tous les éléments

à partir du $n^{ième}$; etc. Cette opération ne change pas le caractère distinctif de chaque élément ; l'ensemble caractéristique M du nouveau collectif mathématique est le même que le M de l'ancien. — 2) On peut aussi, pour obtenir la nouvelle suite (e'), faire usage d'un second collectif mathématique H'' dont les éléments soient e''_1, e''_2, e''_3, ... Par exemple, on ne garde l'élément e_λ, pour le mettre dans la nouvelle suite illimitée (e'), que si l'élément correspondant e''_λ a un caractère distinctif faisant partie d'un ensemble partiel déterminé, A'', de l'ensemble caractéristique de H''. — 3) On peut aussi combiner ces préceptes ou les répéter plusieurs fois.

L'axiome II postule que, quelle que soit cette manière de déduire la suite (e') de la suite primordiale (e), pourvu que ce soit par un choix admissible, les rapports des probabilités mathématiques ne soient pas altérés.

Définition 15. — Toute correspondance entre les éléments d'une suite infinie telle que (e), ou de leurs indices, d'une part, et les points d'un spaciom à k dimensions d'autre part, si elle satisfait aux postulats I et II ci-dessus, est dite *arégulière* ou *à caractère fortuit*. On peut dans ces hypothèses formuler la définition du collectif mathématique, esquissée au n° 188 β), comme suit.

Définition 16. — Un collectif mathématique est une suite infinie d'éléments avec une correspondance arégulière entre les caractères distinctifs et les éléments.

192. — Considérons le cas particulier où l'ensemble caractéristique M du collectif mathématique envisagé C est partagé en deux parties seulement, mettons A et B (v. n° 189, définition 11). Il suit du théorème 3 combiné avec le postulat II, que

$$p_A + p_B = p'_A + p'_B = 1,$$

donc, en vertu du postulat II, formule (VIII 9), que

$$p_A = p'_A \text{ et } p_B = p'_B \qquad \text{(VIII 10)}$$

et cela quels que soient les ensembles partiels A et B, pourvu que leur somme soit la totalité de l'ensemble caractéristique M. En d'autres termes, la fonction de fréquence du nouveau collectif

mathématique est égale à celle du collectif primordial. Observons maintenant que toute suite illimitée (e'') que l'on pourrait déduire de (e') par un choix admissible, on pourrait aussi l'obtenir en opérant un choix admissible directement dans la suite primordiale (e). On en conclut que, si la correspondance entre caractère distinctif et élément satisfait aux postulats I et II dans la suite primordiale (e), la même chose a lieu dans (e') ; en d'autres termes : la suite (e') déduite de (e) par un choix admissible constitue à son tour un collectif mathématique, C', dont la fonction de fréquence est identique à celle du collectif primordial C. On peut donc énoncer le

Théorème 4. — Si de la suite (e) d'un collectif mathématique C, on déduit une suite partielle (e') par un choix admissible, c'est-à-dire sans tenir compte des caractères distinctifs des éléments que l'on choisit, on forme un nouveau collectif mathématique, C', dont la diadose ou fonction de fréquence est la même que celle de C.

On peut en particulier déduire du syllepte C deux autres, C_0 et C_1, en mettant dans C_0 tous les éléments e_λ d'indice λ pair et dans C_1 tous ceux d'indice λ impair. Les probabilités dans ces deux nouveaux collectifs mathématiques sont les mêmes que dans C, ce qui du reste ne peut avoir lieu que si la correspondance entre élément et caractère distinctif est arégulière dans le sens de la définition 15 (v. n° 191).

Appliquons ceci aux jeux de hasard. Un des problèmes les plus fameux consiste à trouver un « système de jeu » permettant au joueur, par de savantes combinaisons de mises, d'améliorer ses chances de gain. Le postulat II peut, à ce nouveau point de vue, s'appeler *le postulat de l'impossibilité d'un système de jeu.*

193. — Il pourrait sembler naturel de demander *un exemple d'une correspondance « arégulière »*, c'est-à-dire un précepte permettant de trouver explicitement le caractère distinctif de chaque élément e_λ, connaissant l'indice λ. Or, on peut démontrer que cela est irréalisable. On a les théorèmes suivants.

Théorème 5. — Un collectif mathématique est déterminé par sa diadose ou loi de répartition. Expliciter d'une manière complète et

détaillée la correspondance entre élément et caractère distinctif n'est pas possible par un nombre fini de mots.

Théorème 6. — La correspondance entre élément et caractère distinctif, dans deux collectifs mathématiques C et C' ayant même diadose et par conséquent aussi même ensemble caractéristique M, diffère en une infinité d'endroits. Exprimons cela un peu différemment. Il va bien sans dire que si les deux sylleptes C et C' n'ont pas la même loi de répartition, il existe une infinité d'éléments e_λ n'ayant pas le même caractère distinctif que leur élément homologue respectif e'_λ. Or, la même chose a lieu, d'après le théorème 6, quand les collectifs mathématiques en question ont la même loi de répartition. — Enfin, on déduit des considérations précédentes et de leurs corollaires encore le

Théorème 7. — Il n'existe en général aucune transformation assignable à l'aide d'un nombre fini de mots et permettant de déduire, d'une manière explicite et détaillée, la correspondance qui a lieu dans un collectif mathématique C', entre élément et caractère distinctif, de celle qui a lieu dans un autre collectif mathématique C.

Les postulats I et II, avec les théorèmes explicatifs 1 à 7, déterminent la notion fondamentale de collectif mathématique et en même temps la notion de probabilité mathématique dans le sens le plus moderne. A remarquer qu'on ne peut pas « prouver » l'existence de collectifs mathématiques par une construction analytique, par un exemple concret, comme on peut par exemple prouver l'existence de fonctions analytiques partout continues et ne possédant néanmoins de dérivée en aucun point. Il faut se contenter de l'existence logique, idéelle, des collectifs mathématiques, de leur existence *in abstracto* qui consiste en ceci : on peut soumettre les concepts ainsi définis aux opérations mathématiques sans crainte d'aboutir à des contradictions. Ils sont, comme par exemple la droite géométrique, doués d'existence, mais non de réalité, dans le sens du numéro 133. — Quant à savoir si les définitions sont commodes et pratiquement utilisables, c'est une autre question, sur laquelle nous aurons à revenir.

194. — Avant de procéder à de plus amples développements, je vais intercaler deux remarques mathématiques concernant la diadose ou fonction de fréquence.

a) Il résulte des définitions posées au numéro 189 que la probabilité mathématique p_A de l'apparition du caractère A, dans un collectif mathématique donné C, devient une fonction d'ensemble, bornée, non négative et possédant la propriété additive. Or, il est notablement plus commode d'opérer non avec une fonction d'ensemble, mais avec une fonction de point, c'est-à-dire une fonction ordinaire d'un nombre fini de variables. On peut chercher à représenter la diadose dans le syllepte C, c'est-à-dire la répartition de toutes les valeurs que la probabilité mathématique peut y prendre, par une fonction $p(x_1, x_2, \ldots\ldots, x_k)$ des k coordonnées. Mr de Mises ([102]) y arrive comme suit. Ecrivons la fonction cherchée plus succinctement $p(\overline{x})$, où $\overline{x}$ représente le vecteur de composantes x_1, x_2, $\ldots\ldots$, x_k; puis entendons par $p(\overline{x})$ la valeur que prend p_A pour l'ensemble A dont les points ont comme coordonnées ξ_1, ξ_2, ξ_3, $\ldots\ldots$, ξ_k, pour lesquels

$$\xi_\lambda \leqslant x_\lambda \qquad (\lambda = 1, 2, 3, \ldots\ldots, k) \qquad \text{(VIII 11)}$$

En d'autres termes, $p(\overline{x})$ est la probabilité mathématique pour que la première coordonnée du caractère distinctif ait une valeur *au plus* égale à x_1, la deuxième coordonnée *au plus* la valeur x_2, $\ldots\ldots$, la $k^{\text{ième}}$ coordonnée une valeur *au plus* égale à x_k. Ce $p(\overline{x})$ sera la « loi de répartition » ou « fonction de fréquence » ou « diadose » du collectif mathématique envisagé C, et l'on a le

Théorème 8. — La fonction de fréquence, qui ne peut prendre que des valeurs de l'intervalle clos $< 0; 1 >$, est par rapport à chacune des k variables x_1, x_2, $\ldots\ldots$, x_k monotone, non décroissante et semi-continue à droite. Elle prend la valeur zéro pour $x_1 = x_2 = \ldots\ldots = x_k = -\infty$; elle prend la valeur 1 pour $x_1 = x_2 = \ldots\ldots = x_k = +\infty$. Les trois propriétés

$$0 \leqslant p(\overline{x}) \leqslant 1; \qquad p(-\infty, -\infty, \ldots\ldots, -\infty) = 0;$$
$$p(+\infty, +\infty, \ldots\ldots, +\infty) = 1 \qquad \text{(VIII 12)}$$

([102]) R. E. von Mises, *l. cit.* ([38]), p. 66-67.

se déduisent immédiatement des théorèmes 1 et 3, car, en vertu de (VIII 11), l'ensemble A correspondant à $p(-\infty, \ldots\ldots, -\infty)$ est vide et l'ensemble A correspondant à $p(+\infty, \ldots\ldots, +\infty)$ comprend la totalité de l'ensemble caractéristique M.

b) La *valeur moyenne* de la fonction de fréquence $p(\overline{x})$ est le vecteur $\overline{a}$ dont les composantes sont

$$a_\lambda \equiv \int \xi_\lambda . \, dp(\overline{\xi}) \qquad \text{(VIII 13)}$$
$$(\lambda = 1, 2, \ldots\ldots, k)$$

les intégrales étant prises dans le sens de Stieltjes et étendues à tout le spaciom diacritique, avec l'hypothèse qu'elles ont toutes un sens. On voit que a_λ n'est pas autre chose que la valeur moyenne de x_λ (prise par rapport à p. Voir les numéros 103 et 104).

Si l'on se représente intuitivement p_A sous l'image d'une masse répartie sur les points du spaciom caractéristique M, de telle sorte que tout ensemble A ait une masse $= p_A$, alors $a_1, a_2, \ldots\ldots, a_k$ sont les coordonnées du centre de gravité de la masse ainsi répartie, égale à 1 (voir le numéro 190, remarque 2).

III. — *La déduction de nouveaux collectifs mathématiques.*

195. — De collectifs mathématiques donnés, en déduire d'autres, c'est chose essentielle dans le calcul des probabilités. Nous avons déjà appris à connaître un moyen de réaliser cette opération, c'est *le choix admissible* (v. n° 191). C'est la première des *opérations simples*. Le nouveau collectif mathématique, C', a le même ensemble caractéristique M que le syllepte primordial C, mais d'autres éléments. Ils sont choisis parmi ceux de C sans avoir égard aux caractères distinctifs.

Définition 17. — La deuxième opération simple aboutissant à de nouveaux collectifs mathématiques est *le mélangement*. Elle consiste en ceci : le nouveau collectif mathématique, C', a les mêmes éléments que le syllepte C dont il est déduit, mais les carac-

tères distinctifs x' sont autres. On les déduit des anciens caractères distinctifs, x, par une transformation univoque $x' = f(x)$ dont la fonction inverse ne soit pas univoque. En formule : si $x' = f(x)$ résolue par rapport à x donne $x = \varphi(x')$, il faut que la fonction φ soit plurivoque ou du moins non constamment univalente.

On peut aussi dire : l'ensemble caractéristique M du syllepte primordial C est subdivisé en ensembles partiels φ_λ, et à tous les éléments qui font partie d'un même φ_λ, on fait correspondre un nouveau caractère distinctif x', appartenant à un nouvel ensemble caractéristique M' qui peut ne pas avoir le même nombre de dimensions que M. De cette façon, tous les éléments de C appartenant au même ensemble partiel φ_λ sont en quelque sorte « mélangés », puisqu'on leur attribue à tous le même caractère distinctif nouveau x'_λ; de là le nom de l'opération. En voici trois exemples typiques.

196. — *Premier exemple.* — Supposons que le collectif mathématique primordial, C, ait comme éléments les jets obtenus en nombre illimité en lançant simultanément deux dés. Caractère distinctif : le couple des deux nombres amenés comme points. On a donc deux variables indépendantes, x_1 et x_2, de sorte que $k = 2$ et le spaciom diacritique est bidimensionnel, sp_2. L'ensemble caractéristique M se compose de 36 points dont les coordonnées sont (x_i, x_k) avec $i, k, = 1, 2, \ldots\ldots, 6$. — Par *mélangement*, on en déduit un nouveau collectif mathématique, C' ; il est constitué des mêmes éléments que C, mais chaque élément a comme nouveau caractère distinctif la somme des deux points amenés. Le nouvel ensemble caractéristique, M', est unidimensionnel et comprend les onze nombres 2, 3, 4, ..., 12.

Deuxième exemple. — Les éléments du collectif mathématique primordial, C, sont les droites a_λ tracées dans un même plan par un point fixe P situé à l'intérieur d'un cercle c de ce plan. Caractère distinctif d'un élément : la grandeur de l'angle α_λ formé par cette droite a_λ avec un axe fixe. L'ensemble caractéristique M est donc unidimensionnel et comprend toutes les valeurs entre 0 et 180^0. Appliquons l'opération du mélangement. Le nouveau collectif

mathématique, C', a les mêmes éléments que C, savoir les droites a_λ tracées par le point fixe P dans le plan du cercle c ; mais on attribue à chaque élément un nouveau caractère distinctif : le nombre 1 ou 0, suivant que la corde interceptée par le cercle c sur a_λ est plus courte ou non que le côté du triangle équilatéral inscrit dans c. Le nouvel ensemble caractéristique, M', comprend donc deux points seulement ; c'est une alternative simple, d'après la définition 12 du numéro 190 (v. n° 140, paradoxe de Bertrand).

Troisième exemple. — Le syllepte primordial C a pour éléments les coups de feu tirés contre une cible. Caractère distinctif d'un élément : les coordonnées cartésiennes rectangulaires (x, y) du point touché de la cible. L'ensemble caractéristique M est bidimensionnel et continu, si l'on suppose illimité le nombre des coups tirés. Appliquons l'opération du mélangement. Le nouveau collectif, C', a les mêmes éléments que C, mais on fait correspondre à chaque coup, comme nouveau caractère distinctif, la distance r du point touché au centre de la cible. Le nouvel ensemble caractéristique, M', est donc unidimensionnel et continu.

On *démontre* que si les postulats I et II sont satisfaits dans le collectif primordial C, ils le sont aussi dans C' et que l'on peut calculer la diadose de C', quand on connaît celle de C.

Remarque. — Comme cas limite du mélangement, on pourrait considérer celui où la transformation des caractères distinctifs est uni-univoque. Dans ce cas limite, l'opération ne fournit pas des collectifs mathématiques essentiellement nouveaux.

197. — Mentionnons encore un cas particulier, historiquement intéressant à cause du rôle qu'il a joué dans le développement du calcul des probabilités. Dans la théorie des jeux de hasard, on rencontre presque toujours des collectifs mathématiques dont l'ensemble caractéristique, M, ne comprend qu'un nombre fini s de points et où toutes les probabilités particulières p_A sont égales entre elles.

Définition 18. — On exprime ce fait en disant que *la diadose est uniforme.* Ces p_A égaux entre eux ne sont pas autre chose que les

fameux « cas équipossibles » de la théorie classique (voir les numéros 18, 19, 27). Par l'opération du mélangement (v. n° 195), déduisons d'un collectif mathématique de ce genre, C, un nouveau collectif mathématique, C', qui soit une alternative simple. On peut le faire en partageant en deux groupes les s points de l'ensemble caractéristique de C : on en réunit $f < s$ dans le premier groupe auquel on fait correspondre le nombre 1, et l'on réunit les $(s - f)$ points restants en un second groupe auquel on fait correspondre le nombre 0 (v. n° 190, définition 12). La nouvelle probabilité mathématique devient alors $p' = f : s$; on arrive ainsi au

Théorème 9. — Si d'un collectif mathématique C à diadose uniforme, on déduit par l'opération du mélangement un collectif mathématique C' qui soit une alternative simple, dans ce nouveau collectif, la probabilité mathématique « pour l'arrivée de l'événement E » est égale à $f : s$, rapport du nombre f des « cas favorables à l'événement », au nombre total des « cas également possibles ». Les explications qui précèdent font comprendre le sens des expressions guillemetées qui interviennent ici comme nouveaux termes techniques. Sans la définition préalable des collectifs mathématiques C et C', l'explication du quotient $f : s$ n'aurait pas de sens précis.

198. — *Le départagement* est la troisième « opération simple » permettant de déduire d'un collectif mathématique donné C, dont les éléments sont ceux de la suite (e) (v. n° 188), un nouveau syllepte C'. Ce départagement consiste à fixer certaines composantes des caractères distinctifs. Les caractères distinctifs des éléments restent les mêmes dans C' que dans C, mais C' ne comprend qu'une partie des éléments de C, savoir ceux dont le caractère distinctif appartient à un ensemble partiel bien déterminé, M', de l'ensemble caractéristique M de C. Il faut dès lors supposer que M comprenne plus de deux points, donc que le collectif mathématique primordial C ne soit pas une alternative simple, dans le sens de la définition 12 (v. n° 190). De plus, on suppose que M' comprenne au moins deux ensembles partiels proprement dits, M'_1 et M'_2, à chacun desquels correspondent une infinité d'éléments e_λ. Comme

les caractères distinctifs des éléments ne sont pas modifiés, M' est l'ensemble caractéristique du nouveau syllepte C'.

On démontre encore que si les postulats I et II sont satisfaits dans C, ils le sont également dans C'. Enfin on peut, connaissant la loi de répartition de C, calculer celle de C'.

Exemple. — Envisageons, dans les hypothèses du numéro 33, définition 8, un nombre N d'urnes $U_\lambda(a_\lambda, b_\lambda ; s_\lambda)$ contenant chacune des boules blanches et des boules noires en proportion variable d'une urne à l'autre. Désignons ces proportions par

$$x_1^{(\lambda)} \equiv a_\lambda : b_\lambda \quad (\lambda = 1, 2, 3, \ldots, N) \qquad \text{(VIII 14)}$$

On choisit arbitrairement l'une de ces N urnes et l'on en tire au hasard, s fois de suite, une boule, dans les conditions du numéro 33, définition 8. Une série de s tirages effectués dans ces conditions constitue un élément, e_n, du collectif mathématique C à étudier ($n = 1, 2, 3, \ldots$). Pour chaque série de s tirages, on choisit à nouveau l'urne utilisée, au hasard parmi les N urnes à disposition. Désignons par $x_2^{(n)}$ le nombre des boules blanches apparues à la $n^{\text{ième}}$ série. Attribuons maintenant à e_n, comme caractère distinctif, un couple de deux nombres dont le premier, x_1, est le rapport $x_1^{(\lambda)}$ caractérisant l'urne dans laquelle on fait les s tirages, et dont le second, x_2, est le nombre entier $x_2^{(n)}$ (compris naturellement entre zéro et s) des boules blanches sorties dans la série correspondante. Le spaciom diacritique de C est donc bidimensionnel et l'ensemble caractéristique M se compose de $N.s$ points à deux coordonnées, savoir les couples de nombres $(x_1^{(\lambda)}, x_2^{(n)})$.

Appliquons l'opération du département. Comme ensemble caractéristique M' du nouveau syllepte C', prenons ceux des points (x_1, x_2) dont la seconde coordonnée, x_2, a une valeur constante et fixée d'avance entre zéro et s, mettons $x_2 = 3$, ce qui suppose naturellement $s \geq 3$. La diadose de C' se déduit de celle de C, dans cet exemple, par la règle qui exprime le théorème de Bayes :

$$p'(x_1^{(\lambda)}) = p(x_1^{(\lambda)}, 3) : \sum_\lambda p(x_1^{(\lambda)}, 3). \qquad \text{(VIII 15)}$$

On voit par là que la catégorie de problèmes auxquels on applique habituellement la règle de Bayes rentre dans le cadre général créé par cette nouvelle théorie et qu'il n'est pas indiqué d'en faire un domaine à part complètement séparé du reste.

199. — La quatrième et dernière « opération simple » permettant de passer à de nouveaux collectifs mathématiques est *la combinaison de sylleptes donnés*. Dans les trois premières opérations simples : le choix admissible, le mélangement, le départagement, un seul collectif mathématique C était supposé donné ; pour la quatrième opération, il faut partir de deux collectifs mathématiques (c'est le cas le plus simple), oubien d'un collectif mathématique + une série d'autres, tous donnés. Dans le cas le plus simple, soient donnés les deux sylleptes indépendants C' et C'' ; le premier a comme éléments $e'_1, e'_2, e'_3, \ldots, e'_\lambda, \ldots$, avec les caractères distinctifs $\overline{x}'$ à k' coordonnées, de sorte que son ensemble caractéristique, M', est une multiplicité à k' dimensions (v. n° 188). Le second syllepte, C'', a comme éléments $e''_1, e''_2, e''_3, \ldots, e''_\lambda, \ldots$, avec les caractères distinctifs $\overline{x}''$ à k'' coordonnées, de sorte que son ensemble caractéristique, M'', est un spaciom à k'' dimensions. On déduit les éléments e_λ du nouveau syllepte, C, en accouplant deux à deux ceux des collectifs donnés ; ainsi

$$e_1 \equiv (e'_1, e''_1) ; \qquad e_2 \equiv (e'_2, e''_2) ; \quad \ldots\ldots, \quad e_\lambda \equiv (e'_\lambda, e''_\lambda) ; \ldots$$

Le caractère distinctif $\overline{x}_\lambda$ de chacun des nouveaux éléments e_λ est à $(k' + k'')$ coordonnées, savoir les k' coordonnées du caractère distinctif de e'_λ plus les k'' coordonnées de celui de e''_λ. Le nouvel ensemble caractéristique, M, est donc le produit, dans le sens de la théorie des ensembles, des deux ensembles caractéristiques M' et M'' et constitue une multiplicité à k dimensions, où $k \equiv k' + k''$. Dans ces hypothèses, on démontre ce qui suit.

1° Si les postulats I et II sont satisfaits dans C' et dans C'', ils le sont aussi dans le syllepte dérivé C, et celui-ci est un collectif mathématique.

2° La loi de répartition $p(\overline{x})$ de C se calcule à l'aide des diadoses $p'(\overline{x}')$ de C' et $p''(\overline{x}'')$ de C'' par la formule

$$p(\overline{x}) = p'(\overline{x}') \cdot p''(\overline{x}'') \qquad \text{(VIII 16)}$$

où les coordonnées $x_1, x_2, \ldots, x_k$, de $\bar{x}$ sont : $x_1 = x'_1$; $x_2 = x'_2$; $\ldots$; $x_{k'} = x'_{k'}$; $x_{k'+1} = x''_1$; $x_{k'+2} = x''_2$; $\ldots$; $x_k = x''_{k''}$.

On retrouve ainsi le théorème des probabilités composées (v. n° 33).

Exemple. — Les éléments de C' sont les jets d'un premier dé, ceux de C'' les jets d'un second dé, le caractère distinctif étant chaque fois le nombre de points amené (v. n° 189, exemple). Les éléments, e, du collectif mathématique C sont alors les couples de jets, par exemple le jet simultané des deux dés ; caractère distinctif de e : le couple de nombres, dont les deux coordonnées sont respectivement les points amenés par chacun des deux dés, comme dans le premier exemple du n° 196. Désignons par $p'_1, p'_2, \ldots, p'_6$ les probabilités mathématiques d'amener les points 1, 2, ..., 6 avec le premier dé, et par $p''_1, p''_2, \ldots, p''_6$, celles d'amener respectivement 1, 2, ... , 6 points avec le second dé. Du moment où nous ne supposons pas les dés parfaits, les six nombres p'_λ ne sont pas nécessairement égaux entre eux, mais leur somme est $= 1$ (v. n° 189, remarque 1). La même chose est vraie pour les six nombres p''_λ. On démontre que, dans ces hypothèses,

$$p_{\lambda\mu} = p_\lambda \cdot p_\mu \quad (\lambda, \mu, = 1, 2, 3, 4, 5, 6) \qquad \text{(VIII 17)}$$

est la probabilité mathématique d'amener l'événement composé (λ, μ), c'est-à-dire le nombre de λ points avec le premier dé et simultanément μ points avec le second.

200. — α) A l'aide des quatre « opérations simples » que je viens d'expliquer, on peut définir un nombre illimité d'opérations plus compliquées, en combinant ou composant entre elles les opérations simples, soit 2 à 2 soit 3 à 3 etc., ou en répétant les opérations simples après les avoir composées entre elles (v. n° 45). On voit immédiatement que l'itération d'un choix admissible constitue de nouveau un choix admissible, que l'itération d'un mélangement ou celle d'un départagement donne de nouveau un mélangement ou un départagement, tandis que l'itération de la combinaison simple de deux collectifs mathématiques conduit à des « combinaisons multiples », c'est-à-dire à la combinaison de 3 ou de 4, etc.

ou de n collectifs mathématiques, indépendants ou dépendants les uns des autres. Ceci va nous permettre de préciser et même de définir complètement l'objet du calcul des probabilités.

β) La question de savoir si un jeton est bon ou non (dans le sens du numéro 38), ou la question de savoir si un dé est, ou non, régulier et homogène, si son centre de gravité coïncide ou non avec son centre de figure (v. n° 23, remarque) et, dans le cas de la négative, la question de savoir quelles sont les probabilités mathématiques de l'arrivée des diverses faces, $p(1)$, $p(2)$, ..., $p(6)$, tout cela n'est point un problème du calcul des probabilités, pas plus que mesurer la vitesse d'un corps en mouvement n'est un problème de mécanique rationnelle. De pareilles déterminations sont affaire d'expérience ; on peut les effectuer avec une exactitude plus ou moins grande, en tout cas limitée (voir les expériences de Wolf pour le cas du dé, numéro 189). On n'a un problème de calcul des probabilités que si certaines probabilités mathématiques sont données ou supposées connues et qu'on doive, par le calcul, en déduire d'autres probabilités mathématiques. Par exemple, s'agissant du dé, une fois qu'on a déterminé expérimentalement ou choisi pour $p(1)$, ..., $p(6)$, certains nombres dont la somme $= 1$, on peut se demander quelle est la probabilité mathématique d'amener deux fois de suite le six.

Remarquons l'analogue en mécanique. La Mécanique rationnelle ne répond pas non plus à la question : « Comment tel corps matériel déterminé, que l'on voit actuellement changer de place, se meut-il effectivement en ce moment t_0 ? » ; car c'est là une question de détermination expérimentale, bien que la mécanique céleste puisse vous induire en erreur sur ce point. Par contre, la Mécanique rationnelle apprend, étant données ou supposées connues les positions initiales, les vitesses initiales etc., des points d'un certain système, comment en déduire par le calcul la position et l'état de mouvement du système considéré à un instant t postérieur ou antérieur à t_0.

Quand on veut appliquer le calcul des probabilités à un problème concret, comme d'ailleurs dans toute application d'une théorie à l'expérience, il faut opérer un double passage de la théorie à la réalité physique : 1° Il faut déterminer certaines conditions initia-

les, les circonstances ontologiques du numéro 117, dans notre cas particulier les diadoses des collectifs mathématiques considérés, leurs fonctions de fréquence. 2° Une fois qu'on a effectué tous les développements mathématiques nécessaires et poussé les calculs jusqu'aux résultats numériques, il faut, en partant de ces résultats numériques, tirer des conclusions concernant l'évolution ultérieure des phénomènes. Or, ces deux problèmes-là se trouvent tous les deux en dehors du domaine propre du calcul des probabilités.

γ) On peut résumer les considérations précédentes en formulant comme suit le problème du calcul des probabilités. *L'objet du calcul des probabilités* consiste en ceci : étant donnés des collectifs mathématiques, étudier systématiquement les moyens d'en déduire d'autres, examiner les propriétés de ces collectifs dérivés et trouver comment on peut calculer leurs fonctions de fréquence ou lois de répartition, quand on connaît celle des collectifs donnés.

CHAPITRE IX

LA PROBABILITÉ MATHÉMATIQUE ET L'EXPÉRIENCE

I. — *L'expérience et les deux postulats fondamentaux.*

201. — Les notions de syllepte, de correspondance arégulière, de collectif mathématique etc , définies dans le chapitre précédent, semblent mieux adéquates aux phénomènes naturels que les concepts de la théorie classique des chances introduits au chapitre II. La connexion entre l'univers réel et ces nouveaux concepts est de même nature que la connexion entre les objets matériels et les concepts de la géométrie classique (sphère, cylindre, cône, surface plane, ligne droite, point mathématique etc.), de même nature aussi que la connexion entre les corps pesants et les notions fondamentales de la mécanique rationnelle (corps parfaitement rigide, corps élastique, fluide parfait, pendule mathématique etc.). Dans tous ces cas, l'objet réel a servi de point de départ et l'esprit humain, par voie d'abstraction et par idéalisation, est arrivé peu à peu aux concepts si nets et si clairement définis qui servent de fondement à la science pure.

Je dis qu'à leur tour, les notions de syllepte et surtout de collectif mathématique (v. n° 188, β) sont des idéalisations mathématiquement précises du collectif empirique, objet principal de la stocastique et de la biométrique (v. n° 181).

Pour montrer qu'il en est bien ainsi, je vais d'abord examiner les deux postulats fondamentaux.

202. — Commençons par le postulat I formulé au numéro 189. Il est de plus en plus corroboré par l'expérience. Quand on jette un dé à jouer par exemple, la constance approchée de la fréquence

relative des six faces est un fait expérimental, pourvu que le nombre des jets soit suffisamment grand Si le dé utilisé est géométriquement régulier et physiquement homogène, les fréquences relatives des six faces sont égales entre elles, savoir $= 1/6$. Mais si le dé n'a pas été fabriqué avec grand soin ou s'il est pipé, l'expérience montre encore que les fréquences relatives, quoique différentes entre elles, sont à peu près constantes, et cela d'autant plus que les épreuves sont plus nombreuses. Citons comme exemple les expériences de l'astronome suisse Wolf ([103]). Il nota les résultats de 100 000 épreuves dont chacune consistait à jeter deux dés. De la multitude de ses constatations, extrayons le fait suivant : La fréquence relative des jets où les deux dés amenèrent deux points différents l'un de l'autre fut trouvée

dans une série de 100 jets, égale à 0,88,
» » » » 1 000 » ; » » 0,836,
» » » » 10 000 » , » » 0,8354,
» » » » 100 000 » , » » 0,83533.

D'après la théorie, la probabilité mathématique d'amener, avec deux dés, deux points différents l'un de l'autre est égale à

$$30 : 36 = 5 : 6 = 0{,}83333 \ldots\ldots \text{ (v. n° 23, problème 2).}$$

Pour la nouvelle théorie, les valeurs particulières que prennent les fréquences relatives sont d'importance secondaire. L'essentiel est que ces fréquences relatives, lorsque l'expérience est répétée un nombre suffisant de fois, montrent une répartition telle qu'il soit permis et même plausible d'en conclure, par idéalisation, à l'existence d'une valeur limite que l'on finirait par connaître aussi exactement que l'on voudrait, si l'on pouvait prolonger indéfiniment la suite des épreuves. Le fait que l'expérience conduit à admettre l'existence d'un nombre central autour duquel les fréquences relatives viennent se grouper et se condenser de plus en plus, suffit à justifier le postulat I.

(103) Rudolf Wolf, *Ueber Würfelversuche*; voir Schriften der Naturforsch. Gesellsch. in Bern 1849-1851, Bern 1853; Vierteljahrsschrift der Naturf. Ges. in Zürich, t. 38, Zürich 1893. *Drei Mitteilungen über neue Würfelversuche*, *Ibid.*, t. 26 et 27, Zürich 1881-1883

203. — Dans les applications pratiques à l'étude de phénomènes réels, on n'a jamais affaire à des collectifs mathématiques, infinis par définition, mais à des ensembles statistiques et par conséquent à des suites toujours limitées. Cela entraîne une inexactitude dans les résultats. Par exemple, quand on cherche à évaluer le taux annuel de mortalité, q_x (v. n° 28), les membres du collectif empirique sont les personnes d'âge x, plus exactement celles dont l'âge est compris entre $x - \frac{1}{2}$ et $x + \frac{1}{2}$, habitant une partie déterminée du globe terrestre pendant un laps de temps limité. A chacune de ces personnes, on fait correspondre, comme caractère distinctif, soit le nombre 0 soit le nombre 1, suivant qu'elle atteint l'âge $x + 1$ ou non. L'ensemble caractéristique M se compose alors des deux seuls points 1 et 0 (v. n° 190, définition 12). La liste de répartition permet de déterminer $p(1)$ qui est dans ce cas identique avec le q_x cherché. En fait, on obtient q_x non pas comme limite d'un rapport, mais naturellement comme fraction à numérateur et dénominateur finis, quoique très grands. En calculant avec ce q_x comme s'il était une probabilité mathématique exacte, on arrive à un résultat plus ou moins erroné, et l'inexactitude provient de ce que les ensembles statistiques n'ont pas toutes les propriétés du collectif mathématique.

Or, remarquez que la même chose a lieu dans les applications de la géométrie et de la mécanique rationnelle, par exemple en géodésie, lorsqu'on mesure une ligne de base pour une triangulation. Dans les deux cas, le calcul ne peut s'appliquer à la réalité physique qu'avec un degré plus ou moins grand d'approximation. L'analogie va plus loin encore : dans les deux cas, la Théorie conduit à des propositions permettant d'évaluer l'erreur commise ou du moins de déterminer son ordre de grandeur.

204. — Passons au postulat II (v. n° 191) ; je dis qu'il est justifié également par l'expérience. En effet : 1° On a constaté de tout temps, depuis qu'il existe des jeux de hasard et des statistiques soignées, que la constance approchée de la fréquence relative continue à subsister, et cela avec la même valeur numérique, si,

au lieu d'envisager toutes les épreuves faites, on n'en considère qu'une partie ; par exemple, si l'on ne prend que chaque troisième épreuve ou bien qu'on laisse de côté les 200 premières ou qu'on fasse décider par le sort quelles épreuves doivent être prises en considération, pourvu toutefois que la série partielle envisagée comprenne toujours un nombre suffisamment grand d'observations. En d'autres termes : les chances de deux joueurs dont l'un parie pour le point 1 et l'autre pour le point 3 restent les mêmes si, sans connaître d'avance les résultats des coups, ils conviennent de n'en compter qu'une partie, pourvu qu'elle soit suffisamment nombreuse. Il y a plus : le *rapport* des fréquences relatives n'est pas modifié, si l'on exclut d'emblée tous les éléments qui n'ont pas les caractères distinctifs que l'on veut envisager.

2° Dans *toute* application du calcul des probabilités, l'axiome II est admis. Il revient à l'indépendance des chances. Au jeu de la roulette par exemple, le collectif empirique à étudier comprend tous les coups joués depuis l'existence de la roulette. Or, l'expérience prouve qu'on peut faire abstraction des 10 000 premiers coups, par exemple, et même de toutes les parties jouées avant l'instant t arbitrairement choisi, sans que les fréquences relatives en soient modifiées le moins du monde. Tous les joueurs le savent et l'appliquent. Personne ne pensera sérieusement que les chances de gain, pour les parties dans lesquelles il va s'engager, seront modifiées s'il tient compte des parties jouées la veille ou l'avant-veille ou depuis le début de l'année ou depuis que la roulette existe. Ce fait d'expérience très général, que les chances sont indépendantes les unes des autres et que les fréquences relatives restent invariantes en regard d'un choix admissible (v. n° 191, définition 14), ce fait, dis-je, conduit par idéalisation au postulat II qui en est l'expression mathématique précise.

3° Remarquons encore que dans la pratique de l'assurance sur la vie humaine, notamment dans la construction des tables de mortalité, le postulat II est appliqué très couramment, on peut presque dire inconsciemment : pour déterminer les q_x dont il a besoin, taux annuels de mortalité correspondant aux divers âges x (v. n° 28), l'actuaire se base en fait sur des polices d'assurance déjà échues et sur des décès déjà survenus, donc sur le passé et par conséquent

sur des ensembles partiels ; néanmoins, il considère les résultats de ses investigations comme valables pour les assurés actuels et futurs, pourvu que les observations faites et les groupes envisagés aient été suffisamment nombreux. Donc ici encore, on ne considère que des suites partielles et l'on conclut de la partie au tout, ce qui est bien conforme à l'axiome II.

II. — *Le postulat de la fortuité et les probabilités mathématiques.*

205. — Le postulat II exclut du domaine des probabilités mathématiques proprement dites une catégorie de problèmes traités depuis quelque temps par les méthodes stocastiques, par exemple ceux dont il est question dans les numéros 152 et 153, ou ceux qui traitent de la fréquence relative de tel ou tel chiffre dans les décimales de π ou dans celles d'autres nombres irrationnels. Si l'on adopte les définitions du chapitre VIII, ce genre de problèmes ne rentre pas dans le domaine du calcul des probabilités, parce qu'il s'agit de suites infinies où la correspondance entre l'élément e_λ et son caractère distinctif *n'est pas arégulière* dans le sens de la définition 15 (v. n° 191).

Envisageons une table de racines carrées prolongée aussi loin que l'on voudra, c'est-à-dire contenant les racines carrées des nombres entiers consécutifs, depuis $\sqrt{1} = 1$ jusqu'à $\sqrt{N}$, où N est un nombre naturel aussi grand que l'on voudra Dans cette table, considérons les décimales d'un ordre déterminé r, arbitrairement choisi mais fixe, par exemple $r = 7$. Dans cette suite illimitée de $7^{\text{ièmes}}$ décimales, retenons les n premières, n étant un très grand nombre ; les unes sont paires, les autres impaires ; soit $P_7(n)$ le nombre de celles qui sont paires, $Q_7(n)$ le nombre des impaires, de sorte que $P_7(n) + Q_7(n) = n$. Dans le cas général où l'on envisage les $r^{\text{ièmes}}$ décimales, on a

$$P_r(n) + Q_r(n) = n \qquad \text{(IX 1)}$$

Mr J. Franel (104) a démontré les théorèmes suivants.

a) La fréquence absolue du chiffre l parmi nos $r^{\text{ièmes}}$ décimales, où l représente l'un quelconque des nombres 0, 1, 2, 3, ..., 9, tend vers $\frac{n}{10}$ et par conséquent sa fréquence relative vers 1/10, lorsque n augmente au delà de toute limite. Cela signifie que les divers chiffres 0, 1, 2, 3, ..., 9, sont « en moyenne » répartis uniformément parmi les décimales d'ordre r.

b) La moyenne arithmétique de nos n premières décimales d'ordre r tend vers 4,5 quand n augmente indéfiniment. Or, la somme des dix chiffres de la numération décimale vaut 45, de sorte que 4,5 n'est pas autre chose que la moyenne arithmétique des dix chiffres 0, 1, 2, ..., 9.

c) Le rapport $P_r(n) : n$ tend vers $^1/_2$ quand n augmente indéfiniment. Cela signifie que parmi nos n premières décimales d'ordre r, on rencontre « en moyenne » autant de chiffres pairs que de chiffres impairs, le zéro étant compté comme chiffre pair.

Ces théorèmes prouvent que les choses se passent comme si les divers chiffres, 0, 1, 2, 3, ..., 9, étaient distribués au hasard parmi les $r^{\text{ièmes}}$ décimales des racines carrées des nombres entiers consécutifs ; voir l'analogue au n° 153. Le fait subsiste pour $r=5$, pour $r=6$, pour $r=8$ etc., bref pour les décimales d'un ordre quelconque r suffisamment élevé.

Remarque. — Ce n'est pas là une particularité de la racine carrée. Les mêmes raisonnements restent applicables si, au lieu de $\sqrt{n}$, on envisage $\sqrt[3]{n}$ ou $\sqrt[4]{n}$ ou plus généralement n^{α}, où α représente un nombre positif inférieur à 1, arbitrairement choisi mais fixe. Dans une table des $\alpha^{\text{ièmes}}$ puissances des nombres entiers consécutifs prolongée aussi loin que l'on voudra, où $0<\alpha<1$, la suite illimitée des décimales d'ordre r semble à première vue porter l'empreinte du hasard et constituer un collectif mathématique, malgré la loi rigoureuse qui relie la $r^{\text{ième}}$ décimale de la racine $n^{\text{ième}}$ d'un nombre à ce nombre lui-même. Mais ce n'est

(104) Jérôme Franel, *A propos des tables de logarithmes*, dans la Vierteljahrsschrift der naturforsch. Gesellschaft in Zürich, t. 62, Zürich 1917, fasc. 1 et 2, p. 286-295.

là qu'une ressemblance, car le postulat II n'est pas satisfait, le caractère de fortuité dans le sens de la définition 15 (v. n° 191) fait défaut. La suite illimitée des $r^{\text{ièmes}}$ décimales en question forme donc un syllepte, mais non un collectif mathématique. On le verra plus nettement sur un autre exemple.

206. — Henri Poincaré déjà a reconnu que les mêmes raisonnements sont valables pour les fonctions dérivables quelconques. On démontre même qu'il suffit qu'elles soient continues. Prenons comme exemple la fonction logarithmique. Soit donc une table de logarithmes prolongée aussi loin que l'on voudra, c'est-à-dire contenant les logarithmes de tous les nombres entiers depuis log 1 jusqu'à log N, où N représente un nombre naturel aussi grand que l'on voudra. Envisageons de nouveau, dans cette table, les décimales d'un ordre déterminé quelconque r, arbitrairement choisi mais fixe, par exemple $r = 7$. Dans cette suite illimitée de $7^{\text{ièmes}}$ décimales, retenons les n premières. Les unes sont paires, les autres impaires ; désignons, comme ci-dessus (n° 205), leur nombre respectivement par $P_r(n)$ et par $Q_r(n)$, de sorte que l'égalité (IX 1) ait lieu. Mr Franel ([104]) démontre pour cette suite illimitée trois théorèmes analogues à ceux énumérés ci-dessus (v. n° 205) pour le syllepte des $r^{\text{ièmes}}$ décimales de $\sqrt{n}$; le premier théorème a trait à la fréquence des divers chiffres, 0, 1, 2, 3, ..., 9, le second à la moyenne arithmétique des $r^{\text{ièmes}}$ décimales des n premiers nombres, le troisième à la fréquence relative des chiffres pairs et impairs.

a) Dans la suite des nombres naturels, il existe une infinité de groupes comprenant chacun autant de nombres *consécutifs* qu'on le veut ayant tous cette propriété commune que la $r^{\text{ième}}$ décimale de leur logarithme est $= l$, où l représente l'un des nombres 0, 1, 2, ..., 9.

b) Il existe une infinité de nombres naturels g tels que la moyenne arithmétique des $r^{\text{ièmes}}$ décimales des logarithmes des g premiers nombres entiers diffère aussi peu qu'on le veut de $l + \lambda(r)$, où

$$\lambda(r) \equiv q_r^{-\alpha} \cdot \left\{ 10 \cdot \frac{q_r^{10-l}}{q_r^{10}-1} - \frac{q_r}{q_r-1} \right\} \qquad \text{(IX 2)}$$

où α est arbitrairement choisi entre 0 et 1 et où, pour abréger, on pose

$$q_r \equiv \sqrt[10^r]{10} = 10^{(10-r)} \qquad \text{(IX 3)}$$

Or, quand r est très grand, $\lambda(r)$ a une valeur voisine de 4,5. Donc, à condition que r soit suffisamment grand, c'est-à-dire qu'on envisage les décimales d'un ordre suffisamment élevé, on peut, par un choix approprié du nombre naturel g, faire en sorte que la moyenne arithmétique des g premières décimales d'ordre r diffère aussi peu qu'on le voudra du nombre f, où l'on peut choisir pour f, à volonté, l'un des dix nombres suivants :

$$f = 4\,{}^1/_2,\ 5\,{}^1/_2,\ 6\,{}^1/_2,\ 7\,{}^1/_2,\ \ldots,\ 12\,{}^1/_2,\ 13\,{}^1/_2.$$

c) Un résultat analogue subsiste pour la fréquence des chiffres pairs et des chiffres impairs parmi les décimales d'ordre r dans la table de logarithmes. Soit $P_r(n)$ le nombre des décimales qui sont paires, parmi les n premières décimales d'ordre r. Si les chiffres pairs et les impairs avaient la même fréquence relative parmi les dites décimales, le rapport $P_r(n) : n$ tendrait vers la limite 0,5. Or, chose curieuse, la limite vers laquelle il tend dépend de la manière dont n diverge vers l'infini. Voici les principaux résultats trouvés par Mr Franel.

1° Quand n croît indéfiniment par des puissances de 10, on a

$$\lim_{n \to \infty} \left(P_r(n) : n \right) = \frac{1}{q_r + 1}, \quad \text{donc} \quad < \frac{1}{2}, \qquad \text{(IX 4)}$$

car q_r est toujours légèrement supérieur à 1, comme il ressort de la formule de définition (IX 3);

2° On peut faire croître n vers l'infini par des valeurs telles que $En(10^r . \log n)$ soit toujours pair (v. n° 56, définition 5). Dans ce cas, le rapport $P_r(n) : n$ pourra s'approcher autant qu'on le voudra de

$$1 - \frac{q_r^{1-\alpha}}{q_r + 1} \equiv 1 - \xi, \qquad \text{où} \qquad 0 < \alpha < 1 ; \qquad \text{(IX 5)}$$

3° Si l'on fait croître n vers l'infini par des valeurs telles que $En(10^r . \log n)$ soit toujours impair, le dit rapport devient

$$\frac{P_r(n)}{n} \to \frac{q_r^{1-\alpha}}{q_r + 1} \equiv \xi. \qquad \text{(IX 6)}$$

Or, si r est suffisamment grand, ce nombre ξ diffère peu de $^1/_2$, comme il ressort de la formule de définition (IX 3) ; donc, dans les cas où $\xi < \frac{1}{2}$, la fréquence relative des chiffres pairs, exprimée par le rapport $P_r(n) : n$, est inférieure à $^1/_2$, ce qui signifie que dans les dites décimales d'ordre r, les chiffres pairs sont « en moyenne » moins fréquents que les chiffres impairs.

207. — Les résultats du numéro 206 donnent l'occasion de faire plusieurs remarques. — 1° La fréquence des divers chiffres dans les décimales logarithmiques d'ordre r suit une loi assez compliquée. — 2° Le rapport du nombre d'apparition des chiffres pairs à celui des chiffres impairs, dans les décimales d'ordre r, ne tend pas *dans le sens de l'analyse mathématique* vers une limite déterminée, puisqu'on peut *à volonté*, suivant la manière dont on fait diverger n au delà de toute limite, faire tendre le rapport $P_r(n) : n$ vers ξ ou vers $1 - \xi$. Ceci est nettement en contradiction avec le postulat II. Dès lors, le syllepte constitué par l'ensemble des décimales d'ordre r, dans une table de logarithmes, n'est pas un collectif mathématique et la réserve faite au début du numéro 205 se trouve justifiée. — 3° D'autre part, ξ défini par (IX 6) et par conséquent aussi $1 - \xi$ ont toujours une valeur voisine de $^1/_2$. Le rapport en question oscille donc constamment autour de 0,5 ; tantôt il est supérieur, tantôt inférieur à $^1/_2$, infiniment souvent $= \frac{1}{2}$ quand $n \to \infty$ par la suite naturelle des nombres entiers. Mais quelque grand que l'on prenne n, il existera toujours, dans la suite illimitée des nombres naturels, encore une infinité de nombres n pour lesquels le dit rapport $P_r(n) : n$ tombe en dehors de l'intervalle $< 0{,}49 ; 0{,}51 >$. On peut donc dire que 0,5 est une *limite stocastique* du rapport en question, ou que dans la suite illimitée des $r^{\text{ièmes}}$ décimales logarithmiques, le rapport du nombre d'appa-

rition des chiffres pairs à celui des chiffres impairs tend *stocastiquement* vers 1 (v. n° 59 et n° 60).

208. — On peut construire une infinité d'autres sylleptes dont les éléments sont des nombres ou des chiffres et où la fréquence du zéro ou des autres chiffres se comporte, asymptotiquement, d'une manière systématique qui est en opposition directe avec le postulat II. D'autres fois, il est indifférent pour la solution du problème posé que ce postulat de la fortuité soit satisfait ou non. C'est le cas par exemple de certaines questions traitées par Mr Emile Borel ([105]) et par quelques autres mathématiciens, où il s'agit de la fréquence des divers chiffres dans le développement de certains nombres irrationnels en fraction décimale infinie. Si l'on faisait rentrer sans restriction tous les problèmes de ce genre dans le domaine du calcul des probabilités, la conséquence immédiate en serait l'abandon de la validité générale du théorème des probabilités composées (v. n° 33 et n° 34). Par exemple, la probabilité mathématique de l'apparition de plusieurs zéros consécutifs ne serait plus égale au produit des probabilités mathématiques relatives à chacune de ces décimales Le théorème des probabilités composées pourrait donc tomber en défaut. Or, de tout temps, on a admis que la notion de probabilité impliquait celle de fortuité dans le sens d'imprévisibilité, et comme celle-ci entraîne le théorème de multiplication, on n'aurait plus affaire à des probabilités mathématiques dans le sens classique et courant.

Tout au plus pourrait-on parler dans ce cas de *probabilités mathématiques improprement dites*. Mais il est bien désirable de ne pas ajouter encore une acception nouvelle à ce terme de probabilité qui en a déjà trop ([106]).

(105) Voir Emile Borel, Rendiconti del Circolo matem. di Palermo, t. 27, 1909, p. 247 Il s'agit de problèmes qui reviennent en dernière instance à déterminer la valeur d'une intégrale définie exprimant la mesure d'un ensemble. Ces problèmes peuvent se formuler dans le cadre esquissé au chapitre VIII; il s'agit de « mélangements » dans le sens du n° 195. — Voir aussi Rademacher, Math. Zeitschrift, Berlin, 1918, t. II, p. 306 ; t. III, p. 317 — G. Faber, Math. Annalen, t. 69, 1910, p. 372-443.

(106) F. Hausdorff, par exemple, appelle « probabilité » tout court,

III. — *Les systèmes logiques et l'expérience.*

209. — Reprenons du point de vue philosophique la question de l'applicabilité, au monde réel, des théorèmes du calcul des probabilités. Cette discipline mathématique est-elle purement spéculative ou bien y a-t-il dans la réalité physique quelque chose de corrélatif à ces théorèmes ? Nous répondons qu'il en est du calcul des probabilités comme des autres théories mathématiques. Il faut envisager deux choses distinctes : D'une part, le domaine théorique; là, les propositions peuvent être rigoureusement exactes, parce qu'elles restent, au même titre que les axiomes dont on les déduit, sur le terrain de la logique et de la pure abstraction. D'autre part, le côté application à des domaines concrets ; là, les propositions énoncent ou prétendent énoncer quelque chose sur les phénomènes réels, sur leur évolution ou leurs relations effectives. Il n'est plus dit que les propositions soient encore *rigoureusement* exactes. Elles cessent même de l'être en général. Il en est ainsi en géométrie plane, en stéréométrie, en cinématique, en mécanique etc. Je dis qu'il n'en est pas autrement en stocastique ; le calcul des probabilités peut, au même titre que les autres branches mathématiques, revendiquer d'être un système de propositions cohérentes érigé par déduction logique sur un certain nombre d'axiomes. Je vais expliciter ma pensée ([107]).

Définition 1. — Un système de propositions est dit *logiquement vrai*, s'il y a absence de contradiction entre ses propositions.

La vérité logique d'un système consiste donc en ceci : aucune proposition ni aucun groupe de propositions appartenant au système, ni aucune conséquence s'en déduisant par des opérations logiques, n'est en contradiction avec n'importe quelle proposition du système.

le quotient de la mesure d'un ensemble quelconque par celle d'un ensemble contenant le premier. Voir *Grundzüge der Mengenlehre*, Leipzig, 1914, p. 417.

(107) Voir F.-M. Urban, *l. cit.* (77), p. 44-45.

Le caractère de vérité logique n'appartient donc à une proposition N que par rapport à un groupe donné d'axiomes A ; comme les axiomes peuvent être choisis arbitrairement, parler de « la vérité logique d'une proposition », tout court, n'a pas de sens assignable ; car si une proposition est logiquement vraie par rapport à un groupe d'axiomes A, cela signifie qu'elle appartient nécessairement à l'ensemble des corollaires déduisibles de A ; mais la même proposition est logiquement fausse par rapport à un autre groupe d'axiomes, B, différent de A. Tel est le cas par exemple de la proposition : « La somme des trois angles d'un triangle est égale à deux angles droits » ; elle est juste dans la géométrie euclidienne, mais fausse dans la riemannienne.

210. — Les mathématiques pures sont un ensemble d'affirmations construites d'après le modèle suivant : « Si telle proposition P_1 est vraie d'une chose quelconque N, telle autre proposition P_2 est vraie de la même chose N ». Inutile de chercher à savoir si la première proposition, P_1, est « réellement vraie » ; c'est un axiome. Mr Bertrand Russell va même jusqu'à dire qu'on peut définir les mathématiques pures comme une science « où l'on ne sait pas si ce qu'on dit est vrai ».

Les axiomes eux-mêmes sont indépendants les uns des autres ; du point de vue logique, ils ne sont ni vrais ni faux. Ils sont fortuits les uns par rapport aux autres, parce qu'aucun n'est logiquement déduisible des autres (v. n° 135).

Définition 2. — Une science systématique est un ensemble de signes et de symboles écrits ou parlés, de mots et de propositions que l'on fait dériver d'un groupe d'axiomes conformément aux règles de la logique.

Comme les axiomes dont on part peuvent être choisis tout-à-fait à volonté, pourvu qu'ils ne soient pas en contradiction les uns avec les autres, on voit immédiatement qu'il n'est pas nécessaire que ces signes et symboles soient des images d'objets réels. Les systèmes abstraits sont en dehors de toute expérience sensorielle ; ils ont de l'existence sans avoir de la réalité, dans le sens du numéro 133, définition 4.

Or, il en est ainsi des propositions du calcul des probabilités. Leur vérité logique est complètement indépendante de la question de savoir s'il existe dans le monde physique des choses satisfaisant aux conditions que les dits théorèmes présupposent remplies. Inversement, ni le théorème de Jacques Bernoulli ni celui de Poisson ne permettent de prédire à coup sûr comment les phénomènes se dérouleront effectivement. Poisson ne se rendait pas encore compte de cela, à ce qu'il semble. Lui qui a forgé et introduit le terme de « Loi des grands nombres » [108] croyait non seulement avoir trouvé une généralisation importante du théorème de Jacques Bernoulli au cas où la probabilité mathématique de l'événement étudié varie au cours des épreuves, mais s'imaginait avoir démontré par des raisonnements mathématiques que les effets du hasard *doivent* se compenser, à la longue [109]. Il avait tort. Aucune démonstration logique n'a de prise sur les lois qui régissent l'univers matériel. C'est tout aussi certain que le fait inverse : pour étudier les propriétés d'une sphère ou d'un icosaèdre, il est indifférent que l'on puisse matériellement réaliser les corps géométriques envisagés ou avec quel degré d'approximation cette matérialisation est possible. On peut raisonner juste sur des figures inexactes.

211. — L'intérêt que présentent de pareils systèmes abstraits ne serait pas considérable s'il ne s'en trouvait, parmi eux, dont les signes et les propositions soient, dans le sens de la théorie des ensembles, des *images* d'objets concrets. Ces systèmes, dits *empiriquement vrais*, sont les plus intéressants et les plus importants des systèmes abstraits possibles. La question centrale est donc de savoir si les théorèmes stocastiques sont empiriquement vrais ou s'ils ne sont qu'une occasion de développer des formules et d'exercer la sagacité mathématique.

La mesure d'une probabilité mathématique par une fraction p est basée sur la connaissance de certains faits qui sûrement influencent l'évolution du monde matériel. Mais il se pose tout de

(108) Voir S.-D. Poisson, *Comptes Rendus Acad. Sc. Paris*, t. I, 1835, p. 478.

(109) *Ibidem*, t. II, Paris 1836, p. 377 et suivantes.

même la question suivante : La nouvelle fraction P, déduite de p par des opérations algébriques ou transcendantes plus ou moins compliquées, en tout cas purement abstraites, exprime-t-elle encore quelque chose d'empiriquement vrai ? Un seul événement ne saurait suffire pour résoudre la question ; il faut une longue suite d'épreuves répétées dans des circonstances identiques.

Or, chaque fois qu'on a consulté l'expérience, quand on avait toutes les raisons de croire que les hypothèses faites étaient pratiquement réalisées, on a constaté que les prévisions du calcul étaient corroborées par les faits, en particulier que les écarts de la fréquence relative se répartissaient conformément au théorème de Jacques Bernoulli ; du moins cette confirmation de la théorie par l'expérience a-t-elle été aussi bonne que l'on peut raisonnablement s'y attendre, quand on applique les résultats de déductions purement logiques aux phénomènes inextricablement compliqués du monde réel.

212. — Comment se fait-il que les théorèmes de Jacques Bernoulli et de Poisson se transportent du domaine de la théorie abstraite sur le terrain de la réalité concrète ? La même question, on peut la poser à propos de la géométrie : Comment se fait-il que les théorèmes géométriques, déduits de certains axiomes par des opérations purement logiques, trouvent leur confirmation dans la réalité ? Il n'y a pas d'autre réponse que celle-ci : C'est parce que les axiomes en question sont eux-mêmes empruntés à la réalité, extraits d'elle par idéalisation. Les contours de certains corps matériels sont « pratiquement », c'est-à-dire à très peu de chose près, des lignes droites ; d'autres corps sont « pratiquement » des cubes parfaits, d'autres sont « pratiquement » sphériques etc.

L'analogue se passe en mécanique. Pourquoi les théorèmes de mécanique rationnelle, déduits de certains postulats par des opérations mathématiques, trouvent-ils leur confirmation dans la réalité ? C'est encore parce que les dits concepts fondamentaux sont eux-mêmes modelés sur la réalité : Certains corps se comportent « pratiquement », c'est-à-dire à très peu de chose près, comme s'ils étaient parfaitement rigides, d'autres comme s'ils étaient parfaitement élastiques, d'autres comme pendule mathématique etc.

Qu'en est-il de la stocastique? 1° Les axiomes et les postulats du calcul des probabilités se basent-ils sur la réalité ? 2° Déduit-on les théorèmes de ce calcul par des raisonnements impeccables en partant des dits axiomes et postulats ?

On doit sans contredit répondre affirmativement à cette dernière question. Quant à la première, on remarquera que si l'on base la probabilité mathématique, et avec elle toute la théorie des chances, sur la notion de possibilité, un abîme subsiste entre possibilité et réalité ; et un abîme correspondant s'ouvre entre les phénomènes réels et l'évolution indiquée par les propositions du calcul des probabilités étayé sur cette métaphysique. La vérification expérimentale est alors indispensable, et le métaphysicien pourra s'étonner de voir l'expérience s'accorder si bien avec les prévisions théoriques.

Mais la question change de face dès qu'on base le calcul des probabilités sur la notion de collectif mathématique, comme nous l'avons esquissé dans le chapitre VIII, particulièrement dans les numéros 187 et suivants. Dans ce cas, on voit immédiatement que les concepts fondamentaux de la stocastique sont, eux aussi, extraits de la réalité physique par abstraction et par idéalisation, au même titre qu'en géométrie et en mécanique rationnelle. On doit en conclure que les fréquences relatives observées sont des valeurs approchées des probabilités mathématiques, et de là découle directement l'accord, approché également, entre l'expérience et les prévisions théoriques.

Cette façon de procéder a en outre le très grand avantage de faire d'emblée du calcul des probabilités une discipline mathématique, sans qu'il soit nécessaire de recourir à des considérations d'ordre philosophique ou métaphysique.

APPENDICE

Tables de valeurs de la transcendante de Kramp-Laplace :

$$\Theta(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt.$$

TABLE I

Valeurs de la fonction $\Theta(x)$, de centième en centième,
à sept décimales : de 0,00 à 3,45 avec les différences premières ;
à onze décimales : de 3,46 à 4,00 ; puis de dixième en dixième,
à onze décimales : de 4,0 à 4,8. p. 296-298

TABLE II

Valeurs de la fonction $\Theta(x)$, de millième en millième,
à quatre décimales, pour l'intervalle de 0,000 à 1,509. p. 299-301

Index des noms. p. 302-304

TABLE I

Valeurs de la fonction $\Theta(x)$, à sept et à onze décimales.

$$\Theta(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt.$$

x	$\Theta(x)$	Diff.	x	$\Theta(x)$	Diff.	x	$\Theta(x)$	Diff.
0,00	0,0000 000	112 833	0,40	0,4283 922	95 768	0,80	0,7421 010	59 023
0,01	0,0112 833	112 811	0,41	0,4379 690	94 986	0,81	0,7480 033	58 075
0,02	0,0225 644	112 766	0,42	0,4474 676	94 191	0,82	0,7538 108	57 130
0,03	0,0338 410	112 699	0,43	0,4568 867	93 384	0,83	0,7595 238	56 189
0,04	0,0451 109	112 609	0,44	0,4662 251	92 567	0,84	0,7651 427	55 253
0,05	0,0563 718	112 497	0,45	0,4754 818	91 737	0,85	0,7706 680	54 322
0,06	0,0676 215	112 362	0,46	0,4846 555	90 897	0,86	0,7761 022	53 396
0,07	0,0788 577	112 204	0,47	0,4937 452	90 046	0,87	0,7814 398	52 475
0,08	0,0900 781	112 025	0,48	0,5027 498	89 185	0,88	0,7866 873	51 559
0,09	0,1012 806	111 824	0,49	0,5116 683	88 316	0,89	0,7918 432	50 650
0,10	0,1124 630	111 600	0,50	0,5204 999	87 438	0,90	0,7969 082	49 746
0,11	0,1236 230	111 354	0,51	0,5292 437	86 550	0,91	0,8018 828	48 849
0,12	0,1347 584	111 087	0,52	0,5378 987	85 654	0,92	0,8067 677	47 958
0,13	0,1458 671	110 799	0,53	0,5464 641	84 751	0,93	0,8115 635	47 075
0,14	0,1569 470	110 489	0,54	0,5549 392	83 841	0,94	0,8162 710	46 198
0,15	0,1679 959	110 158	0,55	0,5633 233	82 924	0,95	0,8208 908	45 328
0,16	0,1790 117	109 806	0,56	0,5716 157	82 001	0,96	0,8254 236	44 467
0,17	0,1899 923	109 434	0,57	0,5798 158	81 071	0,97	0,8298 703	43 612
0,18	0,2009 357	109 041	0,58	0,5879 229	80 136	0,98	0,8342 315	42 766
0,19	0,2118 398	108 627	0,59	0,5959 365	79 196	0,99	0,8385 081	41 927
0,20	0,2227 025	108 193	0,60	0,6038 561	78 251	1,00	0,8427 008	41 097
0,21	0,2335 218	107 740	0,61	0,6116 812	77 302	1,01	0,8468 105	40 275
0,22	0,2442 958	107 267	0,62	0,6194 114	76 349	1,02	0,8508 380	39 462
0,23	0,2550 225	106 775	0,63	0,6270 463	75 394	1,03	0,8547 842	38 657
0,24	0,2657 000	106 263	0,64	0,6345 857	74 435	1,04	0,8586 499	37 861
0,25	0,2763 263	105 734	0,65	0,6420 292	73 473	1,05	0,8624 360	37 075
0,26	0,2868 997	105 185	0,66	0,6493 765	72 510	1,06	0,8661 435	36 297
0,27	0,2974 182	104 618	0,67	0,6566 275	71 545	1,07	0,8697 732	35 529
0,28	0,3078 800	104 034	0,68	0,6637 820	70 579	1,08	0,8733 261	34 769
0,29	0,3182 834	103 433	0,69	0,6708 399	69 611	1,09	0,8768 030	34 020
0,30	0,3286 267	102 814	0,70	0,6778 010	68 644	1,10	0,8802 050	33 280
0,31	0,3389 081	102 178	0,71	0,6846 654	67 676	1,11	0,8835 330	32 549
0,32	0,3491 259	101 526	0,72	0,6914 330	66 708	1,12	0,8867 879	31 828
0,33	0,3592 785	100 859	0,73	0,6981 038	65 742	1,13	0,8899 707	31 116
0,34	0,3693 644	100 175	0,74	0,7046 780	64 776	1,14	0,8930 823	30 415
0,35	0,3793 819	99 477	0,75	0,7111 556	63 811	1,15	0,8961 238	29 724
0,36	0,3893 296	98 763	0,76	0,7175 367	62 849	1,16	0,8990 962	29 042
0,37	0,3992 059	98 034	0,77	0,7238 216	61 888	1,17	0,9020 004	28 370
0,38	0,4090 093	97 292	0,78	0,7300 104	60 931	1,18	0,9048 374	27 709
0,39	0,4187 385	96 537	0,79	0,7361 035	59 975	1,19	0,9076 083	27 057

x	$\theta(x)$	Diff.	x	$\theta(x)$	Diff.	x	$\theta(x)$	Diff.
1,20	0,9103 140	26 415	1,70	0,9837 904	6 166	2,20	0,9981 372	872
1,21	0,9129 555	25 784	1,71	0,9844 070	5 958	2,21	0,9982 244	835
1,22	0,9155 339	25 162	1,72	0,9850 028	5 757	2,22	0,9983 079	799
1,23	0,9180 501	24 551	1,73	0,9855 785	5 561	2,23	0,9983 878	764
1,24	0,9205 052	23 949	1,74	0,9861 346	5 371	2,24	0,9984 642	731
1,25	0,9229 001	23 358	1,75	0,9866 717	5 186	2,25	0,9985 373	698
1,26	0,9252 359	22 777	1,76	0,9871 903	5 007	2,26	0,9986 071	668
1,27	0,9275 136	22 206	1,77	0,9876 910	4 832	2,27	0,9986 739	638
1,28	0,9297 342	21 645	1,78	0,9881 742	4 664	2,28	0,9987 377	609
1,29	0,9318 987	21 093	1,79	0,9886 406	4 499	2,29	0,9987 986	582
1,30	0,9340 080	20 552	1,80	0,9890 905	4 340	2,30	0,9988 568	556
1,31	0,9360 632	20 020	1,81	0,9895 245	4 186	2,31	0,9989 124	531
1,32	0,9380 652	19 498	1,82	0,9899 431	4 036	2,32	0,9989 655	507
1,33	0,9400 150	18 987	1,83	0,9903 467	3 892	2,33	0,9990 162	484
1,34	0,9419 137	18 487	1,84	0,9907 359	3 751	2,34	0,9990 646	461
1,35	0,9437 622	17 992	1,85	0,9911 110	3 615	2,35	0,9991 107	441
1,36	0,9455 614	17 510	1,86	0,9914 725	3 482	2,36	0,9991 548	420
1,37	0,9473 124	17 036	1,87	0,9918 207	3 355	2,37	0,9991 968	401
1,38	0,9490 160	16 573	1,88	0,9921 562	3 231	2,38	0,9992 369	382
1,39	0,9506 733	16 118	1,89	0,9924 793	3 111	2,39	0,9992 751	364
1,40	0,9522 851	15 673	1,90	0,9927 904	2 995	2,40	0,9993 115	347
1,41	0,9538 524	15 238	1,91	0,9930 899	2 883	2,41	0,9993 462	331
1,42	0,9553 762	14 811	1,92	0,9933 782	2 775	2,42	0,9993 793	315
1,43	0,9568 573	14 393	1,93	0,9936 557	2 664	2,43	0,9994 108	300
1,44	0,9582 966	13 984	1,94	0,9939 226	2 568	2,44	0,9994 408	286
1,45	0,9596 950	13 585	1,95	0,9941 794	2 469	2,45	0,9994 694	272
1,46	0,9610 535	13 194	1,96	0,9944 263	2 374	2,46	0,9994 966	260
1,47	0,9623 729	12 812	1,97	0,9946 637	2 283	2,47	0,9995 226	246
1,48	0,9636 541	12 483	1,98	0,9948 920	2 194	2,48	0,9995 472	235
1,49	0,9648 979	12 073	1,99	0,9951 114	2 109	2,49	0,9995 707	223
1,50	0,9661 052	11 716	2,00	0,9953 223	2 025	2,50	0,9995 930	213
1,51	0,9672 768	11 367	2,01	0,9955 248	1 947	2,51	0,9996 143	202
1,52	0,9684 135	11 027	2,02	0,9957 195	1 868	2,52	0,9996 345	192
1,53	0,9695 162	10 695	2,03	0,9959 063	1 795	2,53	0,9996 537	183
1,54	0,9705 857	10 370	2,04	0,9960 858	1 733	2,54	0,9996 720	173
1,55	0,9716 227	10 054	2,05	0,9962 581	1 654	2,55	0,9996 893	165
1,56	0,9726 281	9 745	2,06	0,9964 235	1 587	2,56	0,9997 058	157
1,57	0,9736 026	9 444	2,07	0,9965 822	1 522	2,57	0,9997 215	149
1,58	0,9745 470	9 150	2,08	0,9967 344	1 461	2,58	0,9997 364	141
1,59	0,9754 620	8 864	2,09	0,9968 805	1 400	2,59	0,9997 505	135
1,60	0,9763 484	8 585	2,10	0,9970 205	1 343	2,60	0,9997 640	127
1,61	0,9772 069	8 312	2,11	0,9971 548	1 288	2,61	0,9997 767	121
1,62	0,9780 381	8 048	2,12	0,9972 836	1 234	2,62	0,9997 888	115
1,63	0,9788 429	7 789	2,13	0,9974 070	1 183	2,63	0,9998 003	109
1,64	0,9796 218	7 538	2,14	0,9975 253	1 133	2,64	0,9998 112	103
1,65	0,9803 756	7 293	2,15	0,9976 386	1 086	2,65	0,9998 215	98
1,66	0,9811 049	7 055	2,16	0,9977 472	1 039	2,66	0,9998 313	93
1,67	0,9818 104	6 824	2,17	0,9978 511	994	2,67	0,9998 406	88
1,68	0,9824 928	6 598	2,18	0,9979 505	954	2,68	0,9998 494	84
1,69	0,9831 526	6 378	2,19	0,9980 459	913	2,69	0,9998 578	79

x	$\theta(x)$	Diff.	x	$\theta(x)$	Diff.	x	$\theta(x)$
2,70	0,9998 657	75	3,20	0,9999 940	4	3,70	0,9999 9983 285
2,71	0,9998 732	71	3,21	0,9999 944	3	3,71	0,9999 9984 517
2,72	0,9998 802	67	3,22	0,9999 947	4	3,72	0,9999 9985 663
2,73	0,9998 870	63	3,23	0,9999 951	3	3,73	0,9999 9986 726
2,74	0,9998 933	61	3,24	0,9999 954	3	3,74	0,9999 9987 712
2,75	0,9998 994	57	3,25	0,9999 957	3	3,75	0,9999 9988 629
2,76	0,9999 051	54	3,26	0,9999 960	2	3,76	0,9999 9989 477
2,77	0,9999 105	51	3,27	0,9999 962	3	3,77	0,9999 9990 265
2,78	0,9999 156	48	3,28	0,9999 965	2	3,78	0,9999 9990 995
2,79	0,9999 204	46	3,29	0,9999 967	2	3,79	0,9999 9991 672
2,80	0,9999 250	43	3,30	0,9999 969	2	3,80	0,9999 9992 300
2,81	0,9999 293	41	3,31	0,9999 971	2	3,81	0,9999 9992 881
2,82	0,9999 334	38	3,32	0,9999 973	2	3,82	0,9999 9993 421
2,83	0,9999 372	37	3,33	0,9999 975	2	3,83	0,9999 9993 921
2,84	0,9999 409	34	3,34	0,9999 977	1	3,84	0,9999 9994 383
2,85	0,9999 443	33	3,35	0,9999 978	2	3,85	0,9999 9994 812
2,86	0,9999 476	31	3,36	0,9999 980	1	3,86	0,9999 9995 208
2,87	0,9999 507	29	3,37	0,9999 981	1	3,87	0,9999 9995 575
2,88	0,9999 536	27	3,38	0,9999 982	2	3,88	0,9999 9995 915
2,89	0,9999 563	26	3,39	0,9999 984	1	3,89	0,9999 9996 230
2,90	0,9999 589	24	3,40	0,9999 985	1	3,90	0,9999 9996 521
2,91	0,9999 613	23	3,41	0,9999 986	1	3,91	0,9999 9996 790
2,92	0,9999 636	22	3,42	0,9999 987	1	3,92	0,9999 9997 039
2,93	0,9999 658	21	3,43	0,9999 988	1	3,93	0,9999 9997 260
2,94	0,9999 679	19	3,44	0,9999 989	0	3,94	0,9999 9997 482
2,95	0,9999 698	18	3,45	0,9999 989		3,95	0,9999 9997 678
2,96	0,9999 716	17	3,46	0,9999 9900 780		3,96	0,9999 9997 860
2,97	0,9999 733	17	3,47	0,9999 9907 672		3,97	0,9999 9998 028
2,98	0,9999 750	15	3,48	0,9999 9914 101		3,98	0,9999 9998 183
2,99	0,9999 765	14	3,49	0,9999 9920 097		3,99	0,9999 9998 327
3,00	0,9999 779	14	3,50	0,9999 9925 691		4,0	0,9999 9998 458
3,01	0,9999 793	12	3,51	0,9999 9930 905		4,1	0,9999 9999 330
3,02	0,9999 805	12	3,52	0,9999 9935 766		4,2	0,9999 9999 714
3,03	0,9999 817	12	3,53	0,9999 9940 296		4,3	0,9999 9999 881
3,04	0,9999 829	10	3,54	0,9999 9944 519		4,4	0,9999 9999 951
3,05	0,9999 839	10	3,55	0,9999 9948 452		4,5	0,9999 9999 980
3,06	0,9999 849	10	3,56	0,9999 9952 115		4,6	0,9999 9999 992
3,07	0,9999 859	8	3,57	0,9999 9955 527		4,7	0,9999 9999 997
3,08	0,9999 867	9	3,58	0,9999 9958 703		4,8	0,9999 9999 999
3,09	0,9999 876	8	3,59	0,9999 9961 661			
3,10	0,9999 884	7	3,60	0,9999 9964 414			
3,11	0,9999 891	7	3,61	0,9999 9966 975			
3,12	0,9999 898	6	3,62	0,9999 9969 358			
3,13	0,9999 904	6	3,63	0,9999 9971 574			
3,14	0,9999 910	6	3,64	0,9999 9973 636			
3,15	0,9999 916	5	3,65	0,9999 9975 551			
3,16	0,9999 921	5	3,66	0,9999 9977 333			
3,17	0,9999 926	5	3,67	0,9999 9978 990			
3,18	0,9999 931	5	3,68	0,9999 9980 528			
3,19	0,9999 936	4	3,69	0,9999 9981 957			

TABLE II. — *Valeurs de la fonction* $\Theta(x)$, *à quatre décimales.*

x	0	1	2	3	4	5	6	7	8	9
0.00	0.0000	0011	0023	0034	0045	0056	0068	0079	0090	0102
01	0.0113	0124	0135	0147	0158	0169	0181	0192	0203	0214
02	0.0226	0237	0248	0259	0271	0282	0293	0305	0316	0327
03	0.0338	0350	0361	0372	0384	0395	0406	0417	0429	0440
04	0.0451	0462	0474	0485	0496	0507	0519	0530	0541	0552
05	0.0564	0575	0586	0597	0609	0620	0631	0642	0654	0665
06	0.0676	0687	0699	0710	0721	0732	0744	0755	0766	0777
07	0.0789	0800	0811	0822	0833	0845	0856	0867	0878	0890
08	0.0901	0912	0923	0934	0946	0957	0968	0979	0990	1002
09	0.1013	1024	1035	1046	1058	1069	1080	1091	1102	1113
0.10	0.1125	1136	1147	1158	1169	1180	1192	1203	1214	1225
11	0.1236	1247	1259	1270	1281	1292	1303	1314	1325	1336
12	0.1348	1359	1370	1381	1392	1403	1414	1425	1436	1448
13	0.1459	1470	1481	1492	1503	1514	1525	1536	1547	1558
14	0.1569	1581	1592	1603	1614	1625	1636	1647	1658	1669
15	0.1680	1691	1702	1713	1724	1735	1746	1757	1768	1779
16	0.1790	1801	1812	1823	1834	1845	1856	1867	1878	1889
17	0.1900	1911	1922	1933	1944	1955	1966	1977	1988	1998
18	0.2009	2020	2031	2042	2053	2064	2075	2086	2097	2108
19	0.2118	2129	2140	2151	2162	2173	2184	2194	2205	2216
0.20	0.2227	2238	2249	2260	2270	2281	2292	2303	2314	2324
21	0.2335	2346	2357	2368	2378	2389	2400	2411	2421	2432
22	0.2443	2454	2464	2475	2486	2497	2507	2518	2529	2540
23	0.2550	2561	2572	2582	2593	2604	2614	2625	2636	2646
24	0.2657	2668	2678	2689	2700	2710	2721	2731	2742	2753
25	0.2763	2774	2784	2795	2806	2816	2827	2837	2848	2858
26	0.2869	2880	2890	2901	2911	2922	2932	2943	2953	2964
27	0.2974	2985	2995	3006	3016	3027	3037	3047	3058	3068
28	0.3079	3089	3100	3110	3120	3131	3141	3152	3162	3172
29	0.3183	3193	3204	3214	3224	3235	3245	3255	3266	3276
0.30	0.3286	3297	3307	3317	3327	3338	3348	3358	3369	3379
31	0.3389	3399	3410	3420	3430	3440	3450	3461	3471	3481
32	0.3491	3501	3512	3522	3532	3542	3552	3562	3573	3583
33	0.3593	3603	3613	3623	3633	3643	3653	3663	3674	3684
34	0.3694	3704	3714	3724	3734	3744	3754	3764	3774	3784
35	0.3794	3804	3814	3824	3834	3844	3854	3864	3873	3883
36	0.3893	3903	3913	3923	3933	3943	3953	3963	3972	3982
37	0.3992	4002	4012	4022	4031	4041	4051	4061	4071	4080
38	0.4090	4100	4110	4119	4129	4139	4149	4158	4168	4178
39	0.4187	4197	4207	4216	4226	4236	4245	4255	4265	4274
0.40	0.4284	4294	4303	4313	4322	4332	4341	4351	4361	4370
41	0.4380	4389	4399	4408	4418	4427	4437	4446	4456	4465
42	0.4475	4484	4494	4503	4512	4522	4531	4541	4550	4559
43	0.4569	4578	4588	4597	4606	4616	4625	4634	4644	4653
44	0.4662	4672	4681	4690	4699	4709	4718	4727	4736	4746
45	0.4755	4764	4773	4782	4792	4801	4810	4819	4828	4837
46	0.4847	4856	4865	4874	4883	4892	4901	4910	4919	4928
47	0.4937	4946	4956	4965	4974	4983	4992	5001	5010	5019
48	0.5027	5036	5045	5054	5063	5072	5081	5090	5099	5108
49	0.5117	5126	5134	5143	5152	5161	5170	5179	5187	5196
0.50	0.5205	5214	5223	5231	5240	5249	5258	5266	5275	5284
x	0	1	2	3	4	5	6	7	8	9

x	0	1	2	3	4	5	6	7	8	9
0.50	0.5205	5214	5223	5231	5240	5249	5258	5266	5275	5284
51	0.5292	5301	5310	5318	5327	5336	5344	5353	5362	5370
52	0.5379	5388	5396	5405	5413	5422	5430	5439	5448	5456
53	0.5465	5473	5482	5490	5499	5507	5516	5524	5533	5541
54	0.5549	5558	5566	5575	5583	5591	5600	5608	5617	5625
55	0.5633	5642	5650	5658	5667	5675	5683	5691	5700	5708
56	0.5716	5724	5733	5741	5749	5757	5765	5774	5782	5790
57	0.5798	5806	5814	5823	5831	5839	5847	5855	5863	5871
58	0.5879	5887	5895	5903	5911	5919	5927	5935	5943	5951
59	0.5959	5967	5975	5983	5991	5999	6007	6015	6023	6031
0.60	0 6039	6046	6054	6062	6070	6078	6086	6093	6101	6109
61	0.6117	6125	6132	6140	6148	6156	6163	6171	6179	6186
62	0 6194	6202	6209	6217	6225	6232	6240	6248	6255	6263
63	0.6270	6278	6286	6293	6301	6308	6316	6323	6331	6338
64	0.6346	6353	6361	6368	6376	6383	6391	6398	6405	6413
65	0.6420	6428	6435	6442	6450	6457	6464	6472	6479	6486
66	0 6494	6501	6508	6516	6523	6530	6537	6545	6552	6559
67	0 6566	6573	6581	6588	6595	6602	6609	6616	6624	6631
68	0.6638	6645	6652	6659	6666	6673	6680	6687	6694	6701
69	0 6708	6715	6722	6729	6736	6743	6750	6757	6764	6771
0.70	0.6778	6785	6792	6799	6806	6812	6819	6826	6833	6840
71	0 6847	6853	6860	6867	6874	6881	6887	6894	6901	6908
72	0 6914	6921	6928	6934	6941	6948	6954	6961	6968	6974
73	0.6981	6988	6994	7001	7007	7014	7021	7027	7034	7040
74	0.7047	7053	7060	7066	7073	7079	7086	7092	7099	7105
75	0.7112	7118	7124	7131	7137	7144	7150	7156	7163	7169
76	0 7175	7182	7188	7194	7201	7207	7213	7219	7226	7232
77	0.7238	7244	7251	7257	7263	7269	7275	7282	7288	7294
78	0.7300	7306	7312	7318	7325	7331	7337	7343	7349	7355
79	0.7361	7367	7373	7379	7385	7391	7397	7403	7409	7415
0 80	0.7421	7427	7433	7439	7445	7451	7457	7462	7468	7474
81	0 7480	7486	7492	7498	7503	7509	7515	7521	7527	7532
82	0.7538	7544	7550	7555	7561	7567	7572	7578	7584	7590
83	0.7595	7601	7607	7612	7618	7623	7629	7635	7640	7646
84	0 7651	7657	7663	7668	7674	7679	7685	7690	7696	7701
85	0.7707	7712	7718	7723	7729	7734	7739	7745	7750	7756
86	0 7761	7766	7772	7777	7782	7788	7793	7798	7804	7809
87	0.7814	7820	7825	7830	7835	7841	7846	7851	7856	7862
88	0 7867	7872	7877	7882	7888	7893	7898	7903	7908	7913
89	0.7918	7924	7929	7934	7939	7944	7949	7954	7959	7964
0.90	0.7969	7974	7979	7984	7989	7994	7999	8004	8009	8014
91	0.8019	8024	8029	8034	8038	8043	8048	8053	8058	8063
92	0 8068	8073	8077	8082	8087	8092	8097	8101	8106	8111
93	0.8116	8120	8125	8130	8135	8139	8144	8149	8153	8158
94	0 8163	8167	8172	8177	8181	8186	8191	8195	8200	8204
95	0.8209	8213	8218	8223	8227	8232	8236	8241	8245	8250
96	0.8254	8259	8263	8268	8272	8277	8281	8285	8290	8294
97	0.8299	8303	8307	8312	8316	8321	8325	8329	8334	8338
98	0.8342	8347	8351	8355	8360	8364	8368	8372	8377	8381
99	0.8385	8389	8394	8398.	8402	8406	8410	8415	8419	8423
1 00	0.8427	8431	8435	8439	8444	8448	8452	8456	8460	8464
x	0	1	2	3	4	5	6	7	8	9

x	0	1	2	3	4	5	6	7	8	9
1.00	0.8427	8431	8435	8439	8444	8448	8452	8456	8460	8464
01	0.8468	8472	8476	8480	8484	8488	8492	8496	8500	8504
02	0.8508	8512	8516	8520	8524	8528	8532	8536	8540	8544
03	0.8548	8552	8556	8560	8563	8567	8571	8575	8579	8583
04	0.8586	8590	8594	8598	8602	8606	8609	8613	8617	8621
05	0.8624	8628	8632	8636	8639	8643	8647	8650	8654	8658
06	0.8661	8665	8669	8673	8676	8680	8683	8687	8691	8694
07	0 8698	8701	8705	8708	8712	8716	8719	8723	8726	8730
08	0.8733	8737	8740	8744	8747	8751	8754	8758	8761	8765
09	0.8768	8771	8775	8778	8782	8785	8789	8792	8795	8799
1.10	0.8802	8805	8809	8812	8815	8819	8822	8825	8829	8832
11	0.8835	8839	8842	8845	8848	8852	8855	8858	8861	8865
12	0.8868	8871	8874	8878	8881	8884	8887	8890	8893	8897
13	0.8900	8903	8906	8909	8912	8915	8918	8922	8925	8928
14	0.8931	8934	8937	8940	8943	8946	8949	8952	8955	8958
15	0.8961	8964	8967	8970	8973	8976	8979	8982	8985	8988
16	0.8991	8994	8997	9000	9003	9006	9008	9011	9014	9017
17	0.9020	9023	9026	9029	9031	9034	9037	9040	9043	9046
18	0.9048	9051	9054	9057	9060	9062	9065	9068	9071	9073
19	0.9076	9079	9082	9084	9087	9090	9092	9095	9098	9100
1.20	0 9103	9106	9108	9111	9114	9116	9119	9122	9124	9127
21	0.9130	9132	9135	9137	9140	9143	9145	9148	9150	9153
22	0.9155	9158	9160	9163	9165	9168	9171	9173	9176	9178
23	0.9181	9183	9185	9188	9190	9193	9195	9198	9200	9203
24	0.9205	9207	9210	9212	9215	9217	9219	9222	9224	9227
25	0.9229	9231	9234	9236	9238	9241	9243	9245	9248	9250
26	0.9252	9255	9257	9259	9262	9264	9266	9268	9271	9273
27	0.9275	9277	9280	9282	9284	9286	9289	9291	9293	9295
28	0.9297	9300	9302	9304	9306	9308	9310	9313	9315	9317
29	0.9319	9321	9323	9325	9327	9330	9332	9334	9336	9338
1.30	0.9340	9342	9344	9346	9348	9350	9352	9355	9357	9359
31	0.9361	9363	9365	9367	9369	9371	9373	9375	9377	9379
32	0.9381	9383	9385	9387	9389	9390	9392	9394	9396	9398
33	0.9400	9402	9404	9406	9408	9410	9412	9413	9415	9417
34	0.9419	9421	9423	9425	9427	9428	9430	9432	9434	9436
35	0.9438	9439	9441	9443	9445	9447	9448	9450	9452	9454
36	0.9456	9457	9459	9461	9463	9464	9466	9468	9470	9471
37	0.9473	9475	9477	9478	9480	9482	9483	9485	9487	9488
38	0.9490	9492	9494	9495	9497	9499	9500	9502	9503	9505
39	0 9507	9508	9510	9512	9513	9515	9516	9518	9520	9521
1.40	0.9523	9524	9526	9528	9529	9531	9532	9534	9535	9537
41	0.9539	9540	9542	9543	9545	9546	9548	9549	9551	9552
42	0.9554	9555	9557	9558	9560	9561	9563	9564	9566	9567
43	0.9569	9570	9571	9573	9574	9576	9577	9579	9580	9582
44	0.9583	9584	9586	9587	9589	9590	9591	9593	9594	9596
45	0.9597	9598	9600	9601	9602	9604	9605	9607	9608	9609
46	0 9611	9612	9613	9615	9616	9617	9618	9620	9621	9622
47	0.9624	9625	9626	9628	9629	9630	9631	9633	9634	9635
48	0.9637	9638	9639	9640	9642	9643	9644	9645	9647	9648
49	0.9649	9650	9651	9653	9654	9655	9656	9657	9659	9660
1.50	0 9661	9662	9663	9666	9666	9667	9668	9669	9670	9672
x	0	1	2	3	4	5	6	7	8	9

INDEX DES NOMS

Les chiffres indiquent la page.

Alembert, Jean-Lerond d', 7, 23, 25, 39.
Alighieri, Dante, 11.
Aristote, 206.

Bachelier, Louis, 18.
Bayes, Th., 19, 118, 172, 275, 276.
Bénard, H., 226.
Bernoulli, Daniel, 23, 25, 226.
Bernoulli, Jacques, 2, 14, 15, 19, 22, 23, 25, 73, 84, 88, 89, 97, 105 à 151, 166, 181, 183, 206, 211, 213, 214, 217, 221, 230, 242, 292, 293.
Bernoulli, Jean, 2, 25.
Bernoulli, Nicolas, 22.
Berthollet, Claude-Louis, 224.
Bertrand, Joseph, 72, 73, 112, 136, 198, 201, 202, 213, 214, 221, 273.
Bohlmann, G., 29.
Boltzmann, Ludwig, 226, 232, 233, 235, 236, 243 à 245.
Borel, Emile, 29, 70, 125, 231, 289.
Bortkiewicz, Ladislaus von, 246.
Bossuet, Jacques-Bénigne, 6, 28.
Bravais, Auguste, 257.
Broggi, Ugo, 29.
Bruns, Heinrich, 30, 112, 214, 259.
Buffon, George-Louis Leclerc, comte de, 23, 117, 118.

Cardano, Hieronymo, 11.
Carvallo, E., 83.
Carnot, Sadi, 235, 238, 239.
Castelnuovo, Guido, 70, 101, 116, 121.
Cesàro, Ernesto, 199.
Châteaubriand, François-Auguste, vicomte de, 6.
Clausius, Rudolf Julius Emanuel, 239.
Condorcet, Marie-Jean-Antoine-Caritat, marquis de, 2, 23.
Copernic, Nicolas, 12.
Cournot, Antoine-Augustin, 27, 28, 29, 36, 195, 196, 197, 206.
Czuber, Emanuel, 70, 104, 112, 121, 144, 146, 147, 185, 194, 197, 201, 214.

Degen, C. F., 67, 68, 69, 70.
Descartes, René, 7, 164.
Diderot, Denis, 25.
Diophante, 7.
Dirichlet, P. G. Lejeune, 255.
Du Pasquier, L.-Gustave, 40, 220.

Eggenberger, Julius, 115, 118, 119.
Einstein, Albert, 246.
Escott, E. B., 71.
Euclide, 14.
Euler, Léonard, 2, 7, 16, 23, 41, 62, 68, 71, 171, 136, 220.

Faber, G., 289.
Faure, F., 197.
Fechner, Gustav Theodor, 29, 152, 253, 258, 259.
Fermat, Pierre de, 5, 7, 8, 10, 11, 14, 15, 22, 52.
Fourier, Jean-Baptiste-Joseph, baron, 195, 196.
Franel, Jérôme, 285, 286, 287.
Frédéric-le-Grand, 16.

Galilei, Galileo, 12.
Galotti, A., 226.
Galton, Francis, 257.
Gauss, Karl Friedrich, 7, 23, 139, 205, 208 à 211, 234, 253 à 256.
Gay-Lussac, J.-L., 227.
Gerhardt, C. J., 37.
Gouraud, Charles, 3, 7, 8, 11, 14, 17, 23.
Gini, Corrado, 161.
Graunt, John, 20.
Guye, Charles-Eugène, 224, 228, 237.

Halley, Edmond, 21, 22.
Hausdorff, F., 289.
Helm, R., 30.
Hudde, Johann, 21.
Huygens, Christiaan, 15, 19, 21, 22.

Jevons, W. Stanley, 202.

Kant, Immanuel, 169.
Kepler, Johannes, 168.
Köhler, Heinrich Gottlieb, 209, 210.
Kramp, Chrétien, 111, 112.
Kries, Johannes von, 26, 160, 170, 175, 196.

Lämmel, Rudolf, 201.
Lagrange, Joseph-Louis, 7, 14, 16, 23.
Lange, Friedrich Albert, 196.
Langevin, Paul, 246.
Laplace, Pierre-Simon, comte de, 2, 5, 7, 14, 16, 17, 23, 25, 27, 28, 44, 84, 111, 112, 113, 116 à 119, 164, 189, 205, 206, 210, 219, 221.
Le Dantec, Félix-A., 203.
Leibniz, Gottfried Wilhelm, 2, 5, 37, 38.
Lexis, Wilhelm, 222.
Liebmann, H., 93.
Lipps, G. F., 29, 259.
Littré, Maximilien-Paul-Emile, 6.
Loschmidt, 235, 236, 239, 245.
Lourié, Samuel, 26.

Maclaurin, 71.
Markoff, A. A., 93, 112.
Mariotte, l'abbé Edmé, 227.
Maxwell, John Clerk, 226, 232 à 235, 242, 243.
Meinong, A., 185, 187.
Mendel, Georg Johann, 257.
Mentré, F., 27.
Méré, chevalier de, 6, 8, 9, 10, 52.
Milhaud, G., 206.
Mill, John Stuart, 155.
Millikan, Robert Andrews, 226.
Mises, Richard Edler von, 29, 30, 259, 270.
Moivre, Abraham de, 15, 23, 68, 84.
Motel, H. Poterin du, 29.
Morgan, A. de, 111, 155.
Montmort, Pierre-Rémond de, 15.
Morse, Samuel Finley Breese, 66.

Napoléon-le-Grand, 6.
Nernst, Walter, 239.
Neumann, Caspar, 21.
Newton, Isaac, 2, 5, 63, 84, 87, 89, 167, 168.

Opitz, H., 112, 125.

Paciuolo, Luca, 11.
Pascal, Blaise, 1, 5, 7 à 12, 14, 15, 19, 22, 37, 52, 63, 122.
Pearl, Raymond, 247.
Pearson, Karl, 255, 257.
Perrin, Jean, 246.
Petty, William, 20, 21.
Peverone, 11.
Planck, Max, 226.
Poincaré, Henri, 35, 197, 203, 204, 212, 226, 286.
Poisson, Siméon-Denis, 2, 23, 117, 118, 172, 189, 196, 197, 292, 293.
Poterin du Motel, H., 29.
Price, 19.
Pringsheim, Alfred, 25.

Quételet, Lambert-Adolphe-Jacques, 28, 221, 256.
Quiquet, Albert, 203.

Racine, Jean, 6.
Radau, R., 112.

Rademacher, 289.
Russell, Bertrand, 291.

Schooten, Franciscus van, 15.
Spinoza, Baruch, 28.
Stirling, James, 67 à 70, 82, 92, 100, 129, 182.

Tartaglia, Nicolò, 11.

Urban, Friedrich M., 194, 203, 263, 290.

Vastel, L.-G.-C., 22.
Vega, Georg, Freiherr von, 210, 213, 214, 215.
Voltaire, François-Marie Arouet dit, 28.

Weldon, W. F. R., 254.
Weiss, Pierre, 226.
Witt, Jean de, 20, 21.
Wolf, Rudolf, 264, 281.
Wolfskehl, 8.

FIN

IMPRIMERIE BARNÉOUD. LAVAL (MAYENNE). FRANCE.

A LA MÊME LIBRAIRIE

L. GAY

Professeur de chimie-physique à la Faculté des Sciences de Montpellier

LES MATHÉMATIQUES DU CHIMISTE

Un volume in-8 raisin de VIII-208 pages avec figures. . . . . 20 fr.

Ce petit volume qui comble une lacune regrettable rendra de grands services aux élèves des Instituts de chimie qui n'ont aucun ouvrage analogue à leur disposition.

Abbé POTRON

Ancien Professeur aux Universités libres d'Angers et de Lille

EXERCICES DE CALCUL DIFFÉRENTIEL ET INTÉGRAL

Deux volumes in-8 raisin se vendant séparément.

Tome premier : **Résumé théorique et énoncés d'exercices.** Un volume grand in-8 de 332 pages avec nombreuses figures . . . . . 35 fr.

Le tome second est sous presse.

H. OLLIVIER

Professeur à la Faculté des Sciences de Strasbourg

COURS DE PHYSIQUE GÉNÉRALE

A l'usage des candidats au certificat de physique générale, au diplôme d'ingénieur-électricien et à l'agrégation

Troisième édition complètement refondue

Tome premier : **Unités. Gravitation. Électricité et Magnétisme. Ions et Électrons. Symétries.** Un fort volume in-8 raisin de 780 pages avec 421 figures dans le texte et onze planches hors texte en phototypie. . . . . 60 fr.

L'ouvrage complet forme trois volumes. Prix 135 francs, majoration en sus.

W. ROUSE BALL

Fellov and Tutor of Trinity College, Cambridge

RÉCRÉATIONS MATHÉMATIQUES ET PROBLÈMES DES TEMPS ANCIENS ET MODERNES

Traduction de J. FITZ-PATRICK

Tome premier : **Arithmétique. Algèbre et théorie des nombres.** Nouvelle édition. Un volume in-8 de 326 pages avec figures dans le texte . . . . 15 fr.

Tome second : **Questions de géométrie. Questions de mécanique. Questions diverses. Carrés magiques. Problèmes des tracés continus. Équation du troisième degré.** Nouvelle édition. Un volume in-8 de 358 pages avec figures. 15 fr.

Le prix du tome III et dernier (non encore réimprimé) est également de 15 francs.

Extrait du Catalogue des publications de la Librairie Scientifique

J. HERMANN

(Majorations en sus)

BOUTARIC. — **Thermodynamique et Chimie.** 3e éd. de l'ouvrage de P. Duhem. 1927 . . . . . 50 fr. »»

MACH. — **La mécanique.** Trad. Bertrand. Nouvelle édition, avec portraits gravés . . . . . 45 fr. »»

CAMPBELL (M. R.). — **La Théorie électrique moderne,** in-8 de 464 pages, 1919 . . . . . 36 fr. »»

CAMPBELL (M. R.). — 1er *Supplément.* **Théorie quantique des Spectres. La relativité.** Trad. Corvisy, 1924, 240 pages . . . . . 18 fr. »»

CAMPBELL (M. R.). — 2e *Supplément.* **La Structure del'Atome** 1925 20 fr. »»

BECQUEREL (Jean). — **Cours de Physique à l'usage des Etudiants et des Ingénieurs.** 2 volumes 1925-27 . . . . . 70 fr. »»

BOUTROUX (P.). — **Les principes de l'analyse mathématique.** Exposé historique et critique. 2 volumes, 1913-1919 . . . . . 100 fr. »»

FABRY (E.). — **Problèmes d'Analyse Mathématique.** 1912, gr. in-8 30 fr. »»

TANNERY (J.). — **Introduction à la Théorie des fonctions d'une variable.** 2e édition en 2 volumes, 1904-1911 . . . . . 75 fr. »»

ASTON (F.-W.). Professeur au Collège de la Trinité à Cambridge, Lauréat du Prix Nobel. — **Les Isotopes.** Traduit par Mlle S. Veil, avec Préface de M. G. Urbain, membre de l'Institut, Paris, 1923, gr. in-8 de xix-164 pages. . . . . 20 fr. »»

FABRY (E.). — **Nouveau Traité de Mathématiques générales,** 4e éd., 2 volumes, 1925 . . . . . 70 fr. »»

FABRY (E.). — **Problèmes et Exercices de Mathématiques générales.** 2e éd., 1913 (second tirage 1920) . . . . . 30 fr. »»

HADAMARD (J.). — **Leçons sur le calcul des variations.** T. Ier, 1911. 36 fr. »»

FABRY (E.). — **Théorie des Séries à termes constants. Applications aux calculs numériques,** 1912 . . . . . 13 fr. »»

PERRY (J.), trad. Davaux. — **Mécanique appliquée.** Tome Ier. **L'Energie mécanique** (avec 205 fig.), 1913. — Tome II. **Constructions déformables et machines en mouvement.** 2 vol . . . . . 60 fr. »»

Sir J.-J. THOMSON F. R. S. — **Les Rayons d'Electricité Positive et leur application aux analyses chimiques.** Trad. Fric et Corvisy. Paris, 1923, in-8 de x-224 pages, avec 9 planches hors texte . . . . . 20 fr. »»

GOURSAT (Ed.). — **Leçons sur l'intégration des équations aux dérivées partielles du premier ordre.** 2e éd., *entièrement refondue,* 1921, 460 pages. 50 fr. »»

CARTAN (E.). — **Leçons sur les invariants intégraux.** 1922 . . 30 fr. »»

GOURSAT (Ed.). — **Leçons sur le problème de Pfaff** . . . . 40 fr. »»

LEMOINE (Paul). — **Traité pratique de Géologie.** 2e édition, 1922, 64 pl. 40 fr. »»

NERNST. — **Chimie Générale.** 2e éd., 2 vol., 1922-24 . . . . . 120 fr. »»

ANDOYER (H.). — **Cours d'astronomie,** Première partie : **Astronomie théorique,** 3e édition, entièrement refondue. — 1 vol. gr. in-8 de 455 pages avec figures. Paris, 1923 . . . . . 45 fr. »»

ANDOYER (H.) et LAMBERT. — **Astronomie pratique** (Seconde partie). Gr. in-8, 1924, 320 pages, 10 planches . . . . . 40 fr. »»
La Troisième partie (**Astronomie physique**) est *sous presse.*

BOREL (Em.). — **Éléments de la Théorie des Probabilités.** 3e édion, 1925. 25 fr. »»

HADAMARD (J.). — **Cours d'Analyse de l'Ecole Polytechnique.** 2 vol. in-8. (Vol. I *en souscription*) . . . . . 60 fr. »»

LE CHATELIER. — **Leçons sur le carbone, la combustion, les bois chimiques,** 2e édit., 1926 . . . . . 30 fr. »»

ROUSE BALL. — **Récréations mathématique et problèmes des temps anciens et modernes,** 3 volumes in-8 . . . . . *chaque* 15 fr. »»

GIBBS (J.-W.). — **Mécanique, Statistique.** Trad. Rossignol. 1926. 30 fr. »»

SWARTS (Fr.). — **Chimie inorganique.** 4e édition, 1926 . . . . 60 fr. »»

LAVAL. — IMPRIMERIE BARNÉOUD.

UNIV. OF CALIFORNIA
WITHDRAWN

14 DAY USE
RETURN TO DESK FROM WHICH BORROWED
ASTRONOMY, MATHEMATICS-
STATISTICS LIBRARY
This book is due on the last date stamped below, or
on the date to which renewed.
Renewed books are subject to immediate recall.

OCT 26 1966

AUG 17 1967

LD 21–50m-4,'63
(D6471s10)476

General Library
University of California
Berkeley

QA273.D7
-
-
C037445997
U.C. BERKELEY LIBRARIES
C037445997
QA 273
D7
UNIV. OF CALIFORNIA
WITHDRAWN